



schools online curriculum content initiative

MOBILE APPLICATION FOR LANGUAGE LEARNING MALL Research Project Report

MAY 2009



Acknowledgements

The Le@rning Federation gratefully acknowledges the work of all the schools and teachers who contributed to this research project.

The teachers actively involved in the project were:

Kerry Ambrose, Deryn Mansell, Karl Krause, Sue Roberts, Norma Ellis, Sally Letcher, Di Hutchings, Ann Louise Moir, Tim Rodgers, Deanne Saunders, Thea Maguire.

The MALL project team

The Le@rning Federation: Leanne Robertson, Gayl O'Connor, Robyn Holdsworth, Celia Mitchell

Learnosity Pty Ltd: Mark Lynch, Gavin Cooney

Educational consultant: Caryl Oliver (mobile learning solutions)

Subject-matter expert: Russell Darnley OAM

Quantitative Data Analyst: Sandy Muspratt (Honorary Research Fellow, School of Education, The University of Queensland)

Disclaimer

The material contained in this MALL Research Project Report is for general information purposes only. Any use of the material is at your own risk. To the extent permitted by law, Curriculum Corporation will not be liable for any loss or damage suffered as a result of any party relying upon this MALL Research Project Report.

Contents

1	EXECUTIVE SUMMARY	1
2	INTRODUCTION.....	4
2.1	Mobile Application for Language Learning project.....	4
2.2	The Le@rning Federation	4
2.3	Key research questions	4
2.4	Research methods.....	5
3	THE PROJECT	6
3.1	Indonesian language	6
3.1.1	Selecting Indonesian.....	6
3.1.2	Creating the conversations	6
3.2	The technology	7
3.2.1	Computer and forum	7
3.2.2	Communication tools	7
3.3	Selecting the schools.....	7
3.4	Conduct of the project.....	8
3.4.1	Preparation	8
3.4.2	Process	9
3.4.3	Support for teachers and students.....	10
3.4.4	Project questionnaires	10
4	FINDINGS	11
4.1	Quantitative analysis.....	11
4.1.1	Pre- and post-testing.....	11
4.1.2	Student questionnaire questions	19
4.1.3	Conclusions from quantitative analysis	20
4.2	Qualitative analysis	22
4.2.1	Students	22
4.2.2	Teachers	23
4.3	Focus group outcomes	24
4.3.1	Is the Mobile Application for Language Learning approach a viable option for teaching and learning of languages?.....	24
4.3.2	Is this approach a viable option for inclusion in the overall blend of teaching languages to students?.....	25
4.3.3	Has this approach improved the language skills of the participating students?	26
4.3.4	Is there potential for a broader based implementation of the technology being used in the trial?.....	27
5	RECOMMENDATIONS.....	29
5.1	General.....	29
5.2	Mobile phones.....	29
5.3	Timing	29
5.4	Indonesian language	29
5.5	Content.....	30
5.6	School infrastructure.....	30
6	APPENDIXES	31
6.1	Appendix 1: Additional quantitative analysis	31
6.2	Appendix 2: How it works	39
6.3	Appendix 3: Questionnaire instrument content and participant group	41

1 Executive summary

The Mobile Application for Language Learning project investigated the implementation of mobile technology as a means by which students may interact with a computer to conduct conversations in Indonesian.

Indonesian was selected because it is a language that is taught in a broad range of schools and one which has existing digital curriculum resources from The Learning Federation to assist with teaching. Four conversation themes were prepared in which students:

- 1 discussed what they saw in a series of photos depicting life in Indonesia
- 2 gave directions to a taxi driver to reach certain landmarks in Jakarta following the map provided
- 3 explained the menu to a customer from the perspective of a waiter/waitress
- 4 provided information to a traveller on certain resorts from the perspective of a travel agent helping a client.

The processes of conducting conversations by mobile phone, and then uploading them onto a custom website for marking and commentary, was enabled by a computerised system provided by Learnosity¹.

The MALL project team sought to answer four key research questions:

- 1 Is the Mobile Application for Language Learning approach a viable option for teaching and learning of languages?
- 2 Is this approach a viable option for inclusion in the overall blend of teaching languages to students?
- 3 Has this approach improved the language skills of the participating students?
- 4 Is there potential for a broader-based implementation of the technology being used in the trial?

The student questionnaire responses indicated that students had increased their confidence in using the language and embraced the technology with which the MALL project was delivered. The teacher focus group and interview responses indicated that teachers perceived an improvement in their students' listening and speaking skills in Indonesian.

From schools in Victoria, South Australia and Tasmania, 196 students conducted 1298 conversations with the computer system by means of a mobile phone.

¹ Learnosity (www.learnosity.com) specialises in the use of mLearning in language education.

All students undertook a test conversation at the beginning of the project to determine their level and a similar test at the end of the project to ascertain change or progress. An increase of 11% in the mean score between pre- and post-testing suggests that this project may have had a positive impact on the language skills of the participating students.

The quantitative analysis at the classroom level indicates that statistically positive effects were found for four of the classrooms. A consideration of classroom factors attributing to this effect suggests that the greater the number of corrected tests per student – amounting to increased feedback to the students on their spoken responses to the questions – the greater the anticipated improvement.

While the mean score between pre- and post-testing is one indicator, teachers also reported a marked increase in students' confidence in both listening and attempting to speak the language.

All the participating students already owned their own mobile phone so had no difficulty in taking responsibility for a second device. All students felt that this gave them privacy and freedom to attempt speaking Indonesian that was not possible in a classroom situation. They welcomed the fact that their conversations could be recorded and subsequently reviewed by their teachers.

Teachers also welcomed the facility to listen to each student's conversation as they were better able to identify individual difficulties with the language, vocabulary or understanding.

While students and teachers reported an improvement in listening and speaking skills, there was a concern that the spoken questions were too fast and too colloquial for some students and thus discouraged them early in the project. While the project was devised for students who had been studying Indonesian for more than a year, there were those who had not progressed sufficiently in that time to deal with the phone conversations. There was also a suggestion that, in offering topics with a local Indonesian bent, some students struggled to 'pick' any words that gave them a clue to the question. Other students, however, quickly grasped the topics and were reported to be joking and colloquial in their responses – something they had never demonstrated in class.

There are some recommendations arising out of this project; that:

- 1 this approach to learning languages and other subjects be explored for broader implementation
- 2 a means by which students are given credits with their service providers to support using their own mobile phones be explored and implemented
- 3 timing for projects be discussed with participating schools to ensure maximum opportunity for full implementation and ongoing support

- 4 questions be structured at a greater variety of levels with keywords to assist students in deciphering the rest of the question
- 5 conversation topics be considered that are more familiar to Australian learners
- 6 process documentation be prepared to act as a guide to implement future such projects
- 7 schools be assisted in developing plans to expand access, connectivity and usability of technology in its current form, and with a view to high-speed broadband capability in the future.

2 Introduction

2.1 Mobile Application for Language Learning project

This project investigated the implementation of mobile technology as a means by which students may interact with a computer via mobile phones to conduct conversations in Indonesian. The project targeted 250 (primarily Year 9 and 10) students in 13 schools located in South Australia, Tasmania and Victoria. At the end of the project, 240 students from 11 schools had participated and, while primarily from Years 8, 9 and 10, students in Years 7 and 11 also featured.

The project enabled students to participate in a learning experience using technology with which they were already familiar. The students and teachers reflected on their experiences and provided feedback to the researchers at The Le@rning Federation on how well the technology and the approach of the project fulfilled the following broad learning outcome: *Students learn and consolidate Indonesian language skills using voice-based practice and assessment by means of mobile phones.*

2.2 The Le@rning Federation

The Le@rning Federation is an initiative of all Australian and New Zealand governments to produce digital curriculum resources (K–12). Learning objects are digital curriculum resources that combine graphics, text, video, audio, animation and interactive tools to engage and motivate student learning across the range of learning areas and levels of schooling. This project is an extension of these activities by virtue of incorporating the ubiquity of mobile phones, an emergent technology in education, within the teaching and learning process.

2.3 Key research questions

- 1 Is the Mobile Application for Language Learning approach a viable option for teaching and learning of languages?
- 2 Is this approach a viable option for inclusion in the overall blend of teaching languages to students?
- 3 Has this approach improved the language skills of the participating students?
- 4 Is there potential for a broader-based implementation of the technology being used in the trial?

2.4 Research methods

The trial was conducted in a total of 11 participating schools located in Victoria, South Australia and Tasmania, with – in the main – one class of students participating from each school. The students had access to a mobile phone continuously during the study period, and teachers accessed the online marking site throughout the study period. Teachers nominated themselves to participate in the study. They were required to attend a one-day introductory workshop; distribute phones to their students; have the students complete pre- and post-tests; mark student responses to the pre-, post- and lesson questions; and have the students complete a mid-project and an end-of-project questionnaire.

Students, teachers and school leaders participated in the project, which was designed with a mixed-method approach; that is, one that includes collection and analysis of quantitative and qualitative data. Project participants provided feedback as follows:

- students – formal feedback through mid-project and post-project questionnaires
- teachers – informal feedback throughout the project; focus group and individual interviews at the end of the project; and post-project questionnaire
- school leadership – feedback on school policy on ICT use and mobile devices in schools.

The content of the questionnaires is indicated in Appendix 3.

Student scores from pre-tests and post-tests were compared to indicate where a student had improved. Qualitative analyses of questionnaire and interview responses were conducted by creating data displays in the form of grids, or matrices.

3 The project

Each student was required to call the computer with the mobile phone provided, and respond to a series of questions asked in Indonesian. Students were able to listen to and re-record their answers before submission for marking. The system recorded the number of calls made. On average each student made 6.5 calls that led to a recorded conversation.

3.1 Indonesian language

3.1.1 Selecting Indonesian

Indonesian was selected as the language for this project because it is broadly taught in schools throughout Australia and because the project initially included a text-messaging element which required a language with roman characters. The Le@rning Federation also has a good resource base in the form of digital curriculum resources to support the project and ready access to subject-matter experts.

3.1.2 Creating the conversations

There were four conversation themes:

- 1 discussing what students saw in a series of photos depicting life in Indonesia
- 2 giving directions to a taxi driver to reach certain landmarks in Jakarta following the map provided
- 3 explaining the menu to a customer from the perspective of a waiter/waitress
- 4 providing information to a traveller on certain resorts from the perspective of a travel agent helping a client.

Each topic had a series of approximately 18 questions for each level that were put into conversational sequences of seven questions each. The Le@rning Federation undertook the development of the questions, the translation of those questions into Indonesian and the recording of these by a native Indonesian speaker. The Indonesian speaker also provided model answers to each question.

The total possible score for both the pre-test and post-test was 28. The intention of the pre- and post-tests was to establish the students' language competence and cultural understanding both before and after the project. At the initial teacher workshop, the teachers provided input into the pre- and post-test questions, and some changes to the proposed conversations were made.

The pre-test data also had the specific purpose of placing the student as a Level 1 or Level 2 learner (where the conversation themes remained the same but the conversations were a little more complex) and subsequent questions administered took account of their entry level.

The students were able to select the theme of their conversation, but each conversation was randomly selected by the computer. In general, the students did not have the same 'conversation' twice.

Students recorded 1298 conversations at various levels, of which 945 were completed and corrected; 102 were completed but not corrected; and 251 were started but not completed. There is no data to indicate what percentage of those incomplete tests was a result of technology issues or a decision by the students to withdraw from the conversation. Of the completed and marked conversations, 12% of students discussed photos (theme 1), 22% sought to direct a taxi driver (theme 2), 40% were about food (theme 3) and 26% provided information about certain resorts and hotels (theme 4).

3.2 The technology

3.2.1 Computer and forum

This project was undertaken using a custom-written computer system developed by Learnosity. The system was designed to deliver and record the conversations with the students, upload the sound files to a teacher site for marking, and provide a student site where students could review their mark and listen to model answers.

The Le@rning Federation also provided a forum site where teachers could share and discuss their experiences or issues. The teachers indicated that they were not familiar with using forum sites and this was reflected in their usage of the project site.

3.2.2 Communication tools

Students were provided with mobile phones equipped with Vodafone SIM cards that had been locked down to the extent that they could only phone the computer for their conversations, or their cohorts in the project.

Vodafone coverage was problematic for some regions but students generally were able to access the computer by phone as and when they chose to do so.

All students participating in the project already owned their own mobile phones and used them extensively and so they needed no training on the use of the technology. The students reported no concerns about carrying two mobile phones and overwhelmingly responded that they could use the project phone with ease. Using a mobile phone for school work was a new concept for the majority of the participating students.

3.3 Selecting the schools

Schools were invited to nominate for participation in the project and selection was made according to a combination of geographical spread, and demonstrated levels of support for

Indonesian language learning and innovation in teaching. Table 1 shows the locations of the schools and the number of participating students.

Table 1 The final selection of schools

School	Location	Participating students
A	Victoria	19
B	Victoria	9
C	Victoria	27
D	Victoria	24
E	South Australia	11
F	South Australia	18
G	Victoria	16
H	Tasmania	15
I	South Australia	21
J	Victoria	2
K	Victoria	(*)

(*) School K was essentially a non-participating school, having completed only the pre-test and a minimal number of other tests. No students completed the post-test at this school.

3.4 Conduct of the project

3.4.1 Preparation

The MALL project focused on students viewing stimulus materials (photographs, map, hard copy of a menu, and travel brochure with hotel details); listening (via a mobile phone) to questions in Indonesian about those materials; and responding orally in Indonesian. The students were required to have some prior knowledge of Indonesian, both in listening and interpreting text and culture. While some students were in their first year of learning Indonesian, the majority indicated that they had been engaged with it for at least two years prior to participating in the project.

Three of the topic areas were related to The Le@rning Federation's existing digital curriculum resources from Languages Other Than English 2 (LOTE2), and these resources provided classroom activities during the MALL project: namely: **Photo album**, **Travels in Indonesia: warung** and **Direct a taxi: Jakarta**. Participating teachers were directed to use at least one of these resources both before and during the project.

In general terms, the vocabulary, grammar and cultural understandings employed in the MALL project were drawn from these digital curriculum resources. The key concepts underpinning the digital curriculum resources developed in LOTE2 are that:

- Culture is expressed in language and communication practices.
- Intercultural language learning involves understanding self and other as cultural beings.
- Competent intercultural communication develops in the third place.

Third space learning acknowledges the challenges of intercultural communication in an increasingly globalised world. It seeks to allow the learner to explore and develop the competencies needed to communicate interculturally, seeking to avoid the effects of preconceptions on communication and collaboration.

While this project was not as broad in scope as LOTE2, its key concepts were reflected in questions and desired responses. In such a project it can be challenging to achieve a sense of intercultural communication in the third place. Therefore, not all of the key concepts were equally represented in each set of questions.

The fourth topic area, dealing with accommodation, did not have a counterpart in The Le@rning Federation digital curriculum resources, but is a theme frequently treated in Bahasa Indonesian courses and texts.

The questions about food and accommodation placed the learner in the role of informant. This was a reversal of a common approach in which students are cast as consumers, or visitors to Indonesia who are dependent on informants for their understanding. In reversing the roles, particularly in the area of food, the intention was to foster a greater sense of intercultural communication.

During the development of the questions, the subject-matter expert ensured that the level of formality of the language, grammar, and syntax accorded with contemporary Indonesian language practice. Some students reported early difficulty in understanding the questions but this dramatically improved over the course of the project.

3.4.2 Process

Students completed a pre-test and post-test using the mobile phone. Students had access to their own results and teachers had access to the individual results of students as well as the combined results of the class.

The pre- and post-test questions (and those contained within the body of the project) were developed by a subject-matter expert. The subject-matter expert used professional judgement to provide 'equivalent' questions in the pre- and post-tests: that is, question 1 of the pre-test was judged to be equally as difficult as question 1 of the post-test.

Each pre- and post-test question was designed to be assessed using a sliding scale (of increasing difficulty and complexity) of possible responses according to the following rubric:

- Score 0: No response or completely wrong answer
- Score 1: One word or guessing
- Score 2: Correct answer but very brief, a basic sentence or phrase
- Score 3: Correct answer with more information and more complex sentence
- Score 4: Answer with correct syntax and vocabulary, most complex sentence.

The subject-matter expert provided example responses in a marking guide for the pre- and post-test questions to illustrate these anticipated levels of student response.

3.4.3 Support for teachers and students

Technical support was available to assist teachers and students with any difficulties they had in communicating with the system.

The Le@rning Federation established a forum site where teachers were invited to discuss issues pertaining to the project with The Le@rning Federation support staff and other participating teachers. Generally, this was new technology for the majority of teachers and, while they welcomed the initiative, they did not embrace this forum as they saw it as one too many new things for them to undertake in the overall process. Some teachers indicated that they might participate in a forum site in a longer-term project.

With the mobile phones provided to them, students were capable of calling each other, as well as the computer system. However, few students took advantage of this, although they did network with each other in and out of class. No data was sought as to communication between students using their own mobile phones.

3.4.4 Project questionnaires

Students completed a mid-project questionnaire; and students, teachers and school leaders were asked to complete a post-project questionnaire. Semi-structured individual interviews with selected teachers took place at the conclusion of the project.

The teacher questionnaire focused on feedback on student progress and any issues that arose during the project. The student questionnaires focused on student interaction with the technology.

The school leadership questionnaire focused on informing school policy on ICT use and mobile devices in schools. Only one response to the school leadership questionnaire was received. Accordingly, responses to the school leadership questionnaire are not included in this report.

The questionnaires and semi-structured interviews comprised both open-response and rating-scale questions.

4 Findings

A total of 196 students conducted 1298 conversations with the system, of which 73% were completed by students and marked by teachers. Of those corrected, 27% were pre- and post-tests, and the remainder covered the four conversation topics that were offered.

The quantitative analysis suggests that significant gains occurred from pre- to post-test in four of the participating schools. The qualitative analysis (questionnaires, focus group findings, and semi-structured interviews) indicated that the students increased their confidence in using the language and embraced the technology through which it was delivered.

4.1 Quantitative analysis

4.1.1 Pre- and post-testing

4.1.1.1 General trend

More students completed the pre-test than the post-test, but only 95 completed both, which forms the basis of the pre- and post-test results. Table 2 shows the number of students within each schools who provided a pre-test and a post-test. The number inside the brackets in the third column excludes those who scored zero on the post-test; the number outside the brackets includes those who scored zero on the post-test.

Table 2 Number of students from whom a pre-test and post-test score was obtained

School	Number of students in the data file	Number completing pre-tests and post-tests
A	19	11
B	9	9
C	27	14 (13)
D	24	8
E	11	7
F	18	17
G	16	7 (4)
H	15	11 (9)
I	21	16
J	2	1

The data in Table 2 was analysed using a dependent groups t-test, where the groups are defined as the group of pre-test scores and the group of post-test scores. That is, the groups are 'dependent' because the analysis takes account of the fact that the same students completed the pre-test and the post-test. For the 95 students who completed both the pre-test and post-test – ignoring the fact that they were clustered within classrooms – the difference between the mean pre-test score (44.0) and the mean post-test score (55.7) is statistically significant ($t = 5.45$, $df = 94$, $p < 0.001$), indicating, in general, that students improved during the course of the project.

4.1.1.2 Analysis for initial student level as determined by pre-test

On the basis of the pre-test, students were assigned to one of two levels. If the pre-test score was 60 or less, the student was assigned to Level 1, and if greater than 60 to Level 2. There was one instance where a student was manually assigned to Level 1 by their teacher, as the score was just above 60. During the course of the trial, Level 1 students completed easier material, and Level 2 students were assigned more difficult material. However, both groups completed a common post-test. Mean pre-test and mean post-test scores were separated according to the level to which students were assigned. Table 3 shows the descriptive statistics for the two groups at the pre-test and the post-test, and Table 4 shows a summary of the repeated measures analysis of variance (ANOVA). This analysis is like the dependent groups t-test in that it takes account of the fact that the same students completed the pre-test and the post-test but, in addition, it contains independent groups – the groups defined by level.

Table 3 Means and standard deviations for pre-test and post-test scores for two groups of students

	Pre-test		Post-test	
	<i>Mean</i>	<i>St Dev</i>	<i>Mean</i>	<i>St Dev</i>
Level 1 (n = 65)	33.5	0.16	48.3	0.27
Level 2 (n = 30)	66.8	0.08	71.8	0.22

Table 4 Summary of the ANOVA

Source	df	MS	F	p	η^2
Between subjects	94				
Level	1	3.309	52.55	< 0.001	0.361
Students within Levels	93	0.063			
Within students	95				
Time (pre-test/post-test)	1	0.400	19.112	< 0.001	0.17
Time * Level	1	0.097	4.655	0.034	0.048
Time * Students within levels	93	0.021			
TOTAL	189				

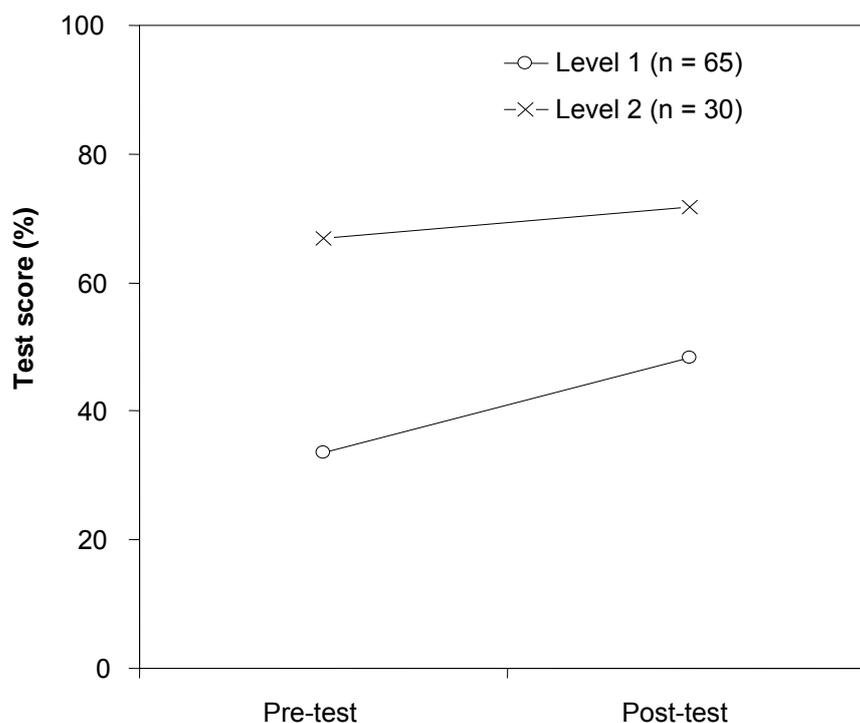
**Figure 1** A graph of the pre-test and post-test means showing the interaction between level and time

Table 4 shows that there is a main effect for level (with Level 2 students doing better than Level 1 students – see Table 3), and a main effect for time (with pre-test scores being lower than post-test scores – see Table 3). However there is a significant interaction between level and time. The interaction is best seen in Figure 1. The interpretation is as follows: even though Level 2 students are doing better than Level 1 students, the Level 1 students have

closed the gap between the two groups at the post-test. This is likely to signify that there is a ceiling effect – Level 2 students who did well at the pre-test have not got much room to improve at the post-test.

4.1.1.3 School-level results

The previous analyses ignore the fact that the students are grouped within schools. It is reasonable to expect that the school a student attends influences their achievements. The following analyses take the grouping of students within schools into account. Three sets of analyses were performed:

- series of dependent groups t-tests
- regression analysis using Markov chain Monte Carlo (MCMC) estimation
- regression analyses and t-test after imputation of missing data.

Series of dependent groups t-test

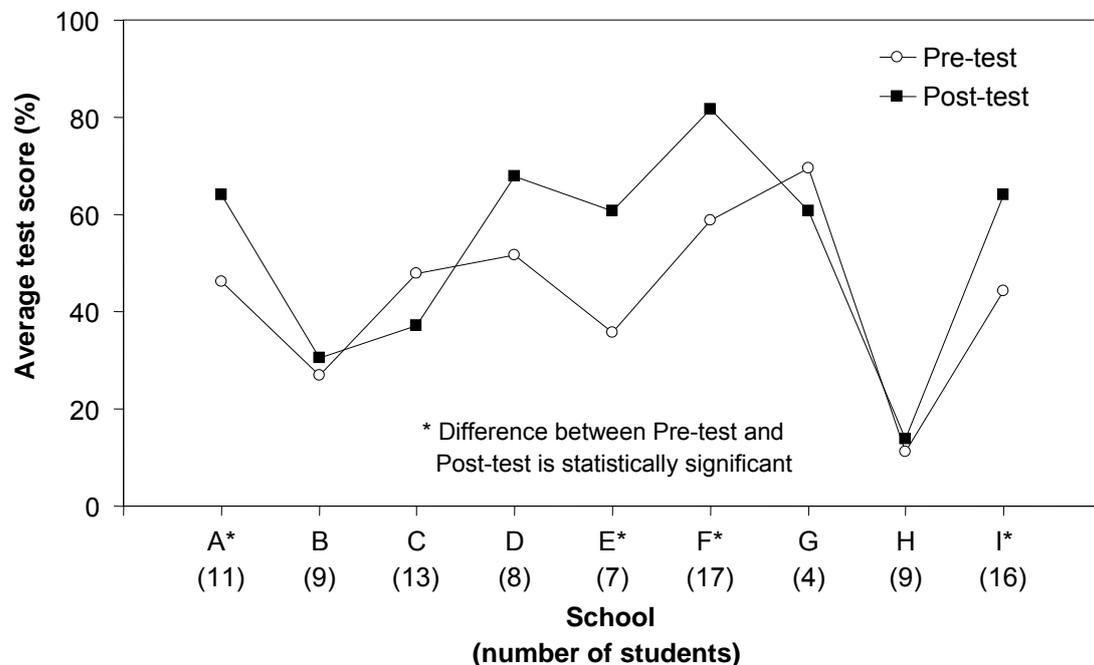
Table 5 shows the results of an analysis that checks for differences between mean pre-test and mean post-test scores for each school, and Figure 2 presents the results graphically. The test is in fact a series of dependent groups t-tests – one for each school (except for School J – a t-test cannot be performed on one individual). In an analysis that involves multiple comparisons, the α -level for each comparison is adjusted so that the family-wise α -level does not exceed 0.05. The right-hand column in Table 5 shows the adjusted α -level for each comparison. If the p-value is less than the α -level, the effect is statistically significant; that is, there is enough evidence to make the claim that the pre-test mean is different from the post-test mean.

For schools A, E, F and I, the differences between the mean pre-test and mean post-test scores are statistically significant. There are two schools that might appear to be going backwards (Schools C and G) in Figure 2, but the differences between mean pre-test and mean post-test scores are not statistically significant.

Table 5 Results of nine dependent groups t-tests (one for each school) testing for differences between pre-test and post-test scores

School	N	Pre-test	Post-test	t	df	p	Holms-adjusted α
A	11	46.1	63.9	4.49*	10	0.001	.0063
B	9	27.0	30.6	< 1			
C	13	47.8	37.1	1.7	12	0.108	.0125
D	8	51.8	67.9	2.6	7	0.037	.0100
E	7	35.7	60.7	4.11*	6	0.006	.0083
F	17	58.8	81.7	7.30*	16	<0.001	.0056
G	4	69.6	60.7	< 1			
H	9	11.1	13.9	< 1			
I	16	44.2	63.9	3.50*	15	0.003	.0071
J	1	60.7	82.1				

* t exceeds critical-t (ie, the difference is statistically significant)

**Figure 2** Pre-test and post-test means for each school

If the analysis also takes into account the level to which a student was assigned, then a number of schools need to be dropped:

- School J has only one student, and was dropped from the previous analysis as well
- Schools B, E and H do not have Level 2 students
- School G has only one Level 1 student (and only four students to begin with); and School D has only two Level 2 students – so for these two schools, detecting any Level effect would be very difficult.

That leaves four schools in the analysis: A, C, F and I. For the students within each school, there is:

- a significant effect for level (in each school, Level 2 students do better than Level 1 students)
- a significant effect for test (in each school, student do better on the post-test than on the pre-test)
- a significant interaction, but only for School C.

Figure 3 shows the pre-test and post-test mean for each level in the four schools. For schools A, F and I, the lines are, in effect, parallel (ie, there is no interaction between level and time). The probable interpretation is that there is insufficient evidence to determine whether the lines deviate. The result for School C, however, is counterintuitive – that the better students do worse at the post-test than the pre-test. It is likely that this result is a consequence of the small numbers – just one or two students doing poorly on the post-test for whatever reason (eg, they had the flu on the day they completed the post-test) is enough to change the overall pattern.

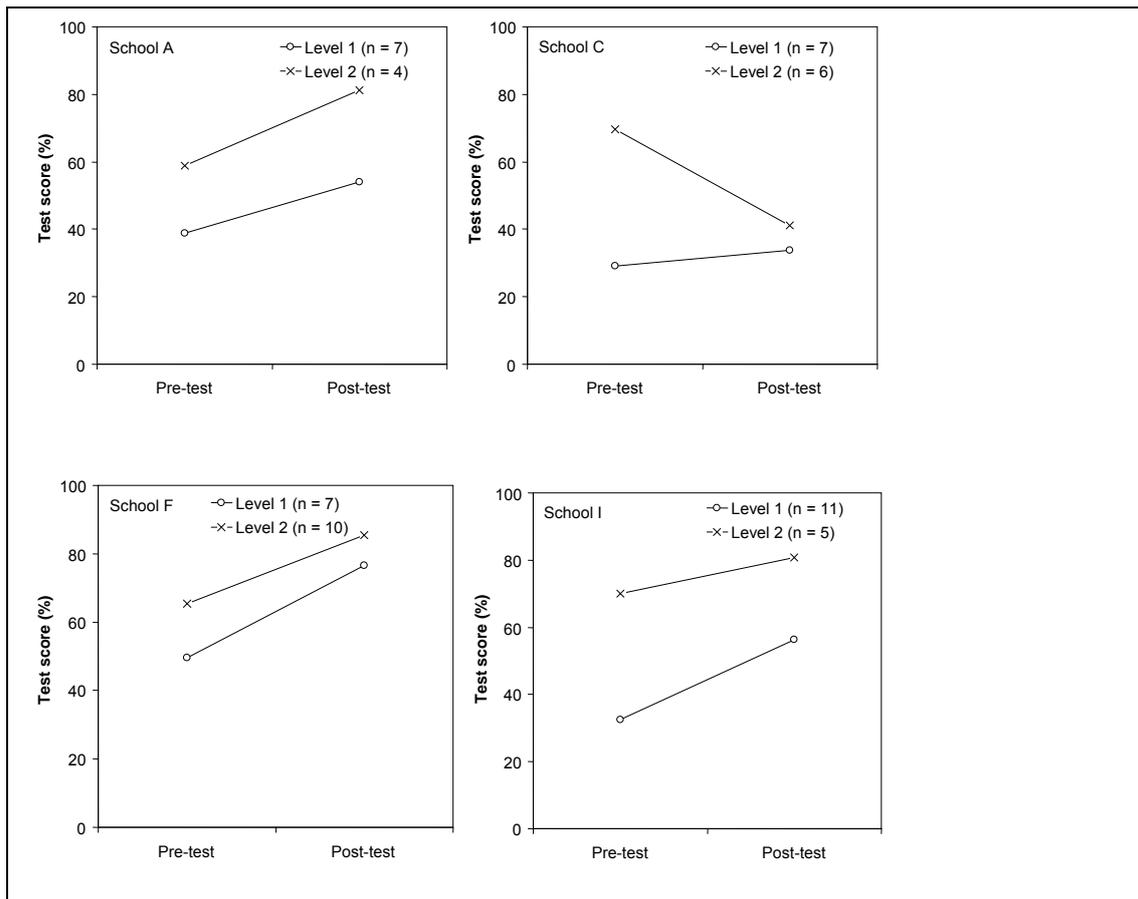


Figure 3 Pre-test and post-test means for each level in Schools A, C, F, I

Regression analysis using Markov chain Monte Carlo (MCMC) estimation

The Markov chain Monte Carlo (MCMC) estimation was also conducted (a single-level analysis within a Bayesian framework with 'school' as the independent variable). The question in this analysis does not focus on the *difference* between pre-test and post-test scores (as in the previous analyses), but the *relationship* between pre-test and post-test scores – the extent to which post-test scores can be predicted from pre-test scores. Details of this methodology are contained in Appendix 1 (6.1.1). The analysis began as a multilevel analysis (see Appendix 1), but because there were too few schools in the data file, the multilevel aspect of the analysis was abandoned. MCMC estimation, though not needed for the following analysis, was carried over from the abandoned multilevel analysis. Table 6 gives the slope of each school's regression line along with a 95% credible interval using this methodology (Appendix 1 (6.1.1) shows the results graphically). Like confidence intervals, if the credible interval spans zero, then the estimate cannot be distinguished from zero. By that criterion, the regression lines for three schools only, are different from zero (Schools A, D and I). But Table 6 shows that, for at least another two schools (B and E), the credible intervals only just span zero. If we are not so strict in application of the criterion, we can say that schools B and E might have a slope that is different from zero. However there is insufficient evidence to be sure and it could very well be that their slopes are not distinguishable from zero. With the

exception of schools G and H, the slopes for the other schools are also in the expected direction (the higher the pre-test score, the higher the post-test score), but more evidence would be needed to be sure that their slopes are different from zero. For the two schools (G and H) that appear to be going backwards, there are not enough students for precise estimates to be made. That is, it is not likely that they are going backwards; however it seems to be the case that too few students are completing both the pre-test and the post-test in these schools. These findings are therefore very similar to those in Table 5.

Table 6 MCMC estimates for the slope of each school's regression line through the scatterplot of pre-test against post-test

School	Slope	St Dev	95% CI
A (11)	1.37	0.47	(0.50, 2.29)
B (9)	0.59	0.40	(-0.19, 1.40)
C (13)	0.25	0.18	(-0.10, 0.62)
D (8)	1.39	0.65	(0.13, 2.67)
E (7)	0.56	0.64	(-0.71, 1.84)
F (17)	0.29	0.38	(-0.46, 1.04)
G (4)	-0.58	0.83	(-2.20, 1.04)
H (9)	-0.56	0.97	(-2.49, 1.33)
I (16)	0.93	0.19	(0.56, 1.30)

Regression analysis and t-test after imputation of missing data

There is missing data in the post-test variable compared to the pre-test variable. Perhaps if the missing data were imputed, the results might be more definitive. In the imputation process, students' pre-test scores, and their responses to the following three questions in the post-project questionnaire, were used:

- By how much did the MALL computer system help you to improve your Indonesian listening skills?
- By how much did the MALL computer system help you to improve your Indonesian speaking skills?
- Overall, by how much did the MALL computer system help you to improve your Indonesian?

Methodology and detailed results are contained in Appendix 1 (6.1.2). The conclusions from the analyses with imputed data are much the same as the conclusions derived from the analyses performed on the original data (for t-tests, student levels and MCMC): that is,

overall, for schools A, E, F and I, the difference between the mean pre-test and mean post-test scores are statistically different.

4.1.2 Student questionnaire questions

The mid-project questionnaire was completed by just 58 students from five schools. Given this small number, it will not be considered further here. However the post-project questionnaire was completed by 91 students (representing the ten schools), giving a response rate of 56%. Table 7 shows the number of students who completed the questionnaire within each school, along with the number of students from the schools who appear somewhere in the data files.

Table 7 Response rates for the post-project questionnaire

School	Number of students in the data file	Number of post-project questionnaires
A	19	7
B	9	6
C	27	20
D	24	5
E	11	8
F	18	14
G	16	6
H	15	13
I	21	11
J	2	1

From the post-project questionnaire, Figure 4 (on page 20) presents the results in terms of students' reported perception of their improvement in various aspects related to learning Indonesian. The error bars are standard errors. The questions asked were:

- By how much did the MALL computer system help you to improve your Indonesian listening skills?
- By how much did the MALL computer system help you to improve your Indonesian speaking skills?
- Overall, by how much did the MALL computer system help you to improve your Indonesian?

- By using the MALL system, how much did your confidence levels increase in being able to listen to and speak Indonesian?

It was found that the greatest means were, in general, for Schools E, F and A.

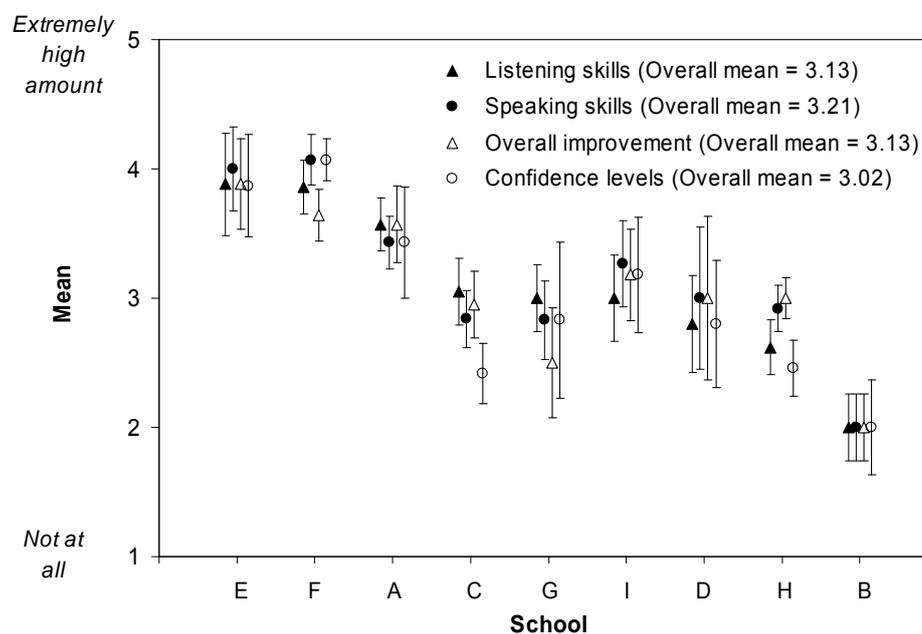


Figure 4 Means for students' perceptions in improvements in listening skills, speaking skills, overall improvement, and improvement in confidence levels

4.1.3 Conclusions from quantitative analysis

From the general trend data and student-level data analysis, it might be claimed that the students did better at the post-test (especially the Level 1 students) because of the work they did during the MALL trial. However, the design will not allow such an unequivocal conclusion. Three obvious counterarguments would be:

- 1 The post-test was easier than the pre-test.
- 2 The students did better at the post-test for no other reason than that they were older and more mature.
- 3 The students did better at the post-test because of the work that the classroom teacher was doing with the students and the improvement had nothing to do with the MALL work.

The first point can be plausibly countered as we know that the pre- and post-test items were constructed by an experienced subject-matter expert who attempted to ensure that question 1 of the pre-test was of equivalent difficulty to question 1 of the post-test and so on. It is also unlikely that a dramatic maturation effect was responsible for the difference in performance,

as the study period was both toward the very end of the academic year and of a short duration (approximately six weeks).

The third point is more difficult to counter. However, despite the difficulties inherent in the experimental design (there were not enough schools to run a full multilevel analyses, and not enough students within many of the schools to allow precise estimates for those schools to be obtained) the analyses that take account of the school allow the opportunity to apply knowledge of what happened in each school during the trial to be applied to the results.

In other words, what else do we know about the schools that might help to explain the differences in student performance? Table 8 shows the number of tests started, completed and corrected by school and per student. As noted above, overall, for schools A, E, F and I, the differences between the mean pre-test and mean post-test scores are statistically different. These schools are annotated with a (#) in Table 8. The students in schools E and F clearly have the greatest number of corrected tests per student, amounting to the most feedback provided to these students by their teachers via the computer system. While the number of corrected tests per student is not as high for School I, the total number of tests attempted was relatively high (11.27 tests per student). It seems reasonable to suggest that the number of tests attempted, and particularly the number of tests corrected by the teacher have a positive impact on student performance. The significant results for School A are somewhat surprising against these criteria but, as can be seen, the number of tests per student is 5.84 and somewhat higher than the next lesser value of 4.93 (School G). The obvious anomaly to this suggested correlation is School C, for which the number of corrected tests per student was high (9.11). The teacher of this class reported some difficulties in terms of student attitudes due to not going on with the subject the following year, and the lateness in the term. This appears to be reflected in the t-test results for School C, where the better students do worse at the post-test than at the pre-test. It is also useful to consider students' perceptions of the various aspects of their Indonesian learning as provided in Figure 4 (on page 20). Students from Schools E, F, A and I reported relatively high means for each aspect. Students from School C felt that their overall performance was about a mean value of 3.13 (lower than Schools E, F and A), with confidence being below the mean (and also below Schools E, F and A).

Table 8 Number of tests started, completed and corrected by school and per student

School	Tests started (and not completed)	Tests completed (but not corrected)	Corrected tests (includes pre-and post-tests)	Test totals	No. students in data file	No. tests any (category per student)	No. corrected tests per student
A #	55	14	42	111	19	5.84	2.64
B	10	2	31	43	9	4.78	3.44
C	64	5	246	315	27	11.67	9.11
D	18	16	42	76	24	3.17	1.75
E #	0	0	124	124	11	11.27	11.27
F #	4	25	196	225	18	12.5	10.89
G	9	0	70	79	16	4.93	4.375
H	18	2	26	46	15	3.07	1.73
I #	47	31	161	239	21	11.38	7.67
J	5	1	7	13	21	0.62	0.33
	251	102	945	1298			

(#) the difference between the pre-test and post-tests means is statistically different.

4.2 Qualitative analysis

4.2.1 Students

The qualitative analysis of student responses is confined to the post-project questionnaire. Of approximately 91 respondents, 71% reported that participating in the project had improved their Indonesian listening skills. Those that felt the project had not improved their listening skills generally attributed this to the questions being spoken too fast, impairing their ability to pick keywords to help them make sense of the questions.

I have learnt many new words and am able to understand them when they're talking now.

Of these 91 students, 76% reported that they had improved their Indonesian speaking skills. There is no clear feedback from students who felt that they had not improved their speaking skills.

I feel more confident in speaking Indonesian and stringing sentences together with flow.

Furthermore, 74% of students reported that the project had helped improve their Indonesian overall. Those that responded negatively indicated that they felt the standard of the language was too high for them so they struggled to understand the questions before they even had to

consider how they might reply. Of particular interest is that 65% indicated that their confidence in using Indonesian had increased.

The MALL project was a really inivative (sic) way to practise our Indonesian skills because it is a way to incorporate something we do all the time, like talking on the phone and practising your Indonesian skills all at once. I loved using the phones and I would definitely love to use the phones again!

When asked if they would like to use a system like MALL again, 25% said never while 48% said only occasionally. Further inquiry highlighted that the major negative lay in poor telephone reception leading to either inability to access the system when desired or to hear clearly when connected. Difficulty in understanding the questions because they were asked too quickly and insufficient 'think' time for responses also featured in the negative responses.

Overall, the students embraced the project and responded positively to the way in which it improved their listening and speaking skills. However, consistent with their generation they were impatient with technology failures and the consequent limitations on their being able to undertake their conversations as and when they wanted to. The percentage of students who reported that the project was fun was 64.

4.2.2 Teachers

Teachers approached the project with enthusiasm and generally reported that they found it easy to incorporate the MALL approach into their teaching plans. The teachers endeavoured to incorporate all four conversation topics into their student preparation, but some focused on just one or two topics as a priority.

Limited access to computers and to the internet did have some negative impact, but not sufficient to deter participation in the project.

... the fact that we tried several times and failed really annoyed them – not to mention a lot of valuable learning time was wasted changing computer rooms etc.

Teachers reported that they found the website for marking student work 'somewhat', to 'extremely', user-friendly.

... it was also way better to hear a student's work without class noise and other distractions. Very easy to pick up more mistakes and common errors than when in a normal class situation.

The participating teachers, bar one, did not have previous familiarity with forum or networking sites and did not embrace the forum site established for this project. The short timeframe for this project may have militated against teachers using the forum site effectively.

I only used it once and then was too busy.

Teachers felt that an approach such as MALL was an effective means of implementing a language course for distance education teachers and teachers who work across a number of

schools. They also responded positively about the effectiveness of the MALL approach in delivering the course material. Despite this positive response, the teachers did comment that where the students found the material too hard they quickly lost interest in the novelty of using mobile phones.

All teachers reported that student engagement was improved by using the MALL approach, as was 'time on task'.

Conversation skills improved, group work was more cohesive and peer teaching was more meaningful. There was greater purpose to learning language.

Teachers generally reported improved student learning outcomes and improved language learning skills as a result of participation in the MALL project.

Although the students, in group conversation, claimed that using the phones didn't help them at all, I believe that some improvement in their confidence occurred.

Overall, the teachers strongly supported the use of mobile phones and online access to their answers for marking, for teaching languages. They indicated some support for the incorporation of The Le@rning Federation digital curriculum resources into their delivery but saw no benefit in the online forum for the teaching process. However they did recognise the benefit of such a support mechanism in the longer term.

4.3 Focus group outcomes

The responses of the focus group participants were organised into data displays in the form of grids, or matrices. This structure assisted in identifying themes and common issues.

4.3.1 ***Is the Mobile Application for Language Learning approach a viable option for teaching and learning of languages?***

Overall, the response to this question was positive. Having a tight lock-down of the phones was seen as important for parents and the schools.

Lock down on phones made parents more comfortable with the program – allowing students to communicate amongst each other was a bonus and an incentive to use the phone more.

Individual teacher responses raised the need for more time to establish the program and increased technical support to ensure computer access for both students and teachers.

Teachers did feel some pressure in listening to and marking the student responses as the volume was significant. However, increasing familiarity with the site and the program helped ease this issue over the course of the project. Some teachers felt that the computer system could assist them more if it had a means to filter those conversations that are ready for marking. Following this suggestion, a filter was added during the course of the project.

Easy. Especially once the option to view only completed sets was available.

Teachers all reported minor problems with the marking and reviewing website and suggestions for improvements to assist teachers with their workflow were incorporated during the course of the project.

Teachers responded positively to the random nature of the questions within the themes, as it encouraged more spontaneous responses from the students. While some teachers felt their students were discouraged if they could not complete the conversation, others felt that their students rose to the challenge.

... allowed teachers to start on a topic, which their students would be most comfortable with and then build their confidence and provide more support for topics they are not familiar with.

The Le@rning Federation forum site was another new element for the teachers who generally did not use it to the extent that they might have done had they been more familiar with this means of communication. They did indicate, however, that it was an aspect that should be retained and that it might have greater usage over a longer time.

Teachers reported that they had endeavoured to use existing digital curriculum resources from The Le@rning Federation, but computer constraints within some schools had limited this activity. Teachers also explored other learning resources to expand student knowledge for the conversations, from Indonesian cookery books to Indonesian websites.

... learning resources downloaded from the net – recipes so that they knew what was in the food to match the questions. Used an Indonesian cooking book ... Looked at U-tube (sic) – but not available at school. Even though not having the phones next year will try more of the conversation stuff.

When asked if teachers saw any impediments to implementing this or a similar initiative more widely, their primary concerns were with computer access and connectivity and the issue of cost in producing and preparing the course material. They were all willing and prepared to incorporate this approach into their teaching and indicated that they would seek professional learning activities that would help them work collaboratively with other teachers.

4.3.2 Is this approach a viable option for inclusion in the overall blend of teaching languages to students?

Absolutely perfectly. It can provide an authentic language experience.

Teachers were overwhelmingly enthusiastic about this approach as they believed that listening and speaking are the main aims for any language student. They also believed that this approach could be viable for other subjects in terms of the form of conversations and the online uploading and marking of student work.

Given the time of year of the project, when teachers believe many students are 'winding down', there was a high level of student engagement where lesson time was used to research

and prepare for the conversations they would have in their own time. All students were reported as improving their listening and speaking skills as well as their confidence.

... blown away by improved confidence. Not afraid to speak in front of others. Weak students; boys in particular.

There was a mixed response to the concept of delivering digital curriculum resources by mobile phone due to the much smaller nature of the device and lack of understanding of the file sizes associated with digital curriculum resources. However, teachers did start to see the opportunities presented by having assignments submitted to a central website and they supported other languages being offered by the same process.

When asked, the teachers were happy to explore ways in which they could integrate mobile phone technology into their teaching.

Would love to get into the podcast by mobile phone area.

Podcasting in particular was seen as a means by which students could collaborate with anyone in the world to improve their language skills and, where the characters are compatible, texting was also seen as an option.

The research showed that increased interaction between students and the system did not automatically mean increased improvement in pre- and post-test scores. While technology as used in this project serves to enhance good teaching practice it is not a substitute for good pedagogy.

4.3.3 Has this approach improved the language skills of the participating students?

When asked to define 'success' at the individual student level, the most common statements were related to improved listening skills and confidence. All teachers reported a significant increase in listening skills and confidence – confidence in this instance meaning students being able to speak in Indonesian in front of their classmates.

All students' confidence increased – they all said they were taking way less time on the phone at the end (weren't holding on pause too long to check possible answers).

At a group level, teachers also reported students helping each other more, improved morale and, in one instance, improved morale of the faculty.

Teachers reported that this approach enabled a level of teacher–student interaction that is not normally always possible and so the project improved the ability of teachers to assess their students directly. They also reported that students took a keener interest in the feedback from the teacher and reflected on their answers by means of the project website. While not every student showed the same degree of improvement, teachers reported that the percentage of success was higher than normal in a class of students undertaking a project.

... all of the students, even those not involved in the MALL project, because they actually listened while I was explaining things to kids and even answered questions which they never did before and wanted to then be involved. Even those that pulled out listened and did it in writing.

By listening to students' recordings teachers were able to make a better assessment of weaknesses in the students speaking Indonesian, with respect to depth of vocabulary, grammar and pronunciation. These are aspects that teachers say they find harder to recognise when working in a class situation with the students.

Teachers reported a marked increase in the confidence of many of their students, exemplified by their starting to use humour in Indonesian. However, they also commented that the questions could have been graded more to cater to a greater range of student ability levels.

Boys improved the most; worked well with everyone but particularly the boys (amount of improvement greater than the girls).

4.3.4 Is there potential for a broader based implementation of the technology being used in the trial?

While teachers responded positively to the potential for broader-based implementation, they were concerned that the computer technology within the schools was structured in such a way as to limit access where and when necessary. Not all project participants had access to computers when they wanted it.

Schools need a more updated computer system – get rid of lots of the filters that hold things up.

Teachers sought more support material and flexibility to use the system under their more direct control in order to select themes and record their own questions. This response is an endorsement of the project as teachers are already considering ways in which to move forward with this technology.

The issue of reception and connectivity was raised as some regions had poor coverage from the provider used in the trial.

Reception – several of my students live in rural areas and had to go to friends' houses or relatives to be able to get reception.

Teachers and students accept that in-school technology has limitations and restrictions. Exploring new and emerging technology as well as overcoming the constraints on campus, presents a challenge that many teachers choose not to accept. In this project, teachers and students were exploring the use of new technology by means of mobile phones with which they were already quite familiar. This familiarity served to flatten the learning curve somewhat. Ubiquitous mobile phones and web-based marking of sound files were manageable but still not always available in the school environment.

While teachers all agreed that this project delivered the learning outcomes and that they would support doing more languages and other subjects in this way, they were concerned that the school infrastructure was not adequate to properly expand the use of technology to support learning.

5 Recommendations

This project clearly demonstrates a positive response to the use of mobile phones for language learning and that there is merit in considering this technology for more broad-based use in teaching languages and other subjects.

With this understanding there are a number of areas that were highlighted in the project that, if improved, would provide a better learning and teaching experience.

5.1 General

This project sparked enthusiasm in both teachers and students by combining ubiquitous technology with teaching practice and formal learning. It generated positive learning outcomes despite some infrastructure issues.

It is recommended that this approach to learning languages and other subjects be explored for broader implementation.

5.2 Mobile phones

Every student who participated in the project, had their own mobile phone. Acquiring mobile phones for the trial was not a major difficulty, but acquiring SIM cards with the level of 'lock-down' required proved to be a significant challenge. While students reported that they had no difficulty with managing two phones this must be taken in the context that they were not restricted from their normal activities with their own phones. Connectivity issues disadvantaged some students as the chosen provider did not have coverage in some areas.

It is recommended that means such as credits paid to service providers, or freecall numbers be explored to enable students to use their own mobile phones.

5.3 Timing

The project was somewhat compromised by having to be completed in a short timeframe at the end of the year, as opposed to allowing longer duration throughout the school year.

It is recommended that timing for projects be discussed with participating schools to ensure the maximum opportunity for full implementation and ongoing support.

5.4 Indonesian language

The project drew in a broader range of standards among the students than anticipated and some were discouraged when they had difficulty understanding the very first interactions they had with the system. Others commented that they had difficulty understanding the questions, which they felt were sometimes too fast and colloquial.

While it is understood that the object of learning languages is to be able to understand locals who do speak fast and not always clearly, there are two recommendations.

It is recommended that:

- *the questions be structured with a greater variety of levels, with keywords to assist students in deciphering the rest of the question*
- *topics, while still in Indonesian, be considered that are more familiar to Australian learners.*

5.5 Content

Because this type of project was being conducted for the first time, considerable effort was required to establish, create and prepare the resources for use on the computer and in the schools. This would be linked to existing resources for use in class teaching during a broader implementation.

It is recommended that process documentation be prepared to act as a guide for implementing such projects in the future.

5.6 School infrastructure

Limited bandwidth, wireless access, computers and computer rooms all represent legacies of pre-digital teaching. This project demonstrates that mobile phones can perform a valuable learning support function if managed and incorporated into teaching practice.

There is the bigger issue of the sheer volume of funding required to bring all schools up to connectivity standards to embrace technology at all levels, but individual schools can make policy and local infrastructure changes that can enhance the way in which technology is used by teachers and students.

It is recommended that schools be assisted in developing plans to expand access, connectivity and usability of technology in its current form, and with a view to high-speed broadband capability in the future.

6 Appendixes

6.1 Appendix 1: Additional quantitative analysis

6.1.1 School level analysis: Markov chain Monte Carlo (MCMC) methodology

Imagine a scatterplot of post-test scores against pre-test scores. A regression line (the line of best fit) can be drawn through the points. The position of the regression line is determined by two coefficients: the intercept and the slope. The intercept is the predicted value for the post-test score when the pre-test score is zero. The slope is the predicted amount by which post-test scores increase for a one unit increase in pre-test scores.

The school level analysis began as a multilevel analysis – one that takes account of the clustering of students within schools. It is reasonable to assume that the school impacts on students' attainments, and moreover, that these effects change from school to school. A consequence of the clustering is that dependencies can arise: students within a school share a common environment, the same teachers, are in direct communication with each other, come from similar neighbourhoods, and so forth. Also, dependencies can arise because students in different classrooms (in effect, different schools) might experience the MALL trial differently. Ignoring these dependencies can lead to spurious significant effects.

Multilevel analyses operate on the student-level and the school-level simultaneously, and return results relevant to the student-levels and the school-level. For the school-level part of the analysis to run properly, multilevel analyses require a reasonable number of schools. For the data at hand, nine schools was not enough. It has been argued² that a Bayesian approach to the analysis can be effective even when the number of schools is small. However, not even a Bayesian analysis was effective on the current data. The Bayesian analysis returned zero variance at the school level, yet it is quite clear that results vary across schools.

The analysis was changed to a single-level analysis in which each school was entered in such a way that 'school' became an effect. The usual way to achieve this is to create a series of dummy variables (one less than the number of schools) to represent the schools in the analysis so that each school can be contrasted against a reference school (the reference school is by default the first school in the list). The planned analysis included the level to which students were assigned (Level 1 or Level 2 according to the results on the pre-test), and asked whether level influenced the slope of the regression line. However, given the difficulties encountered when level was introduced to the analysis (see 4.1.1.2), that analysis was soon abandoned. Instead, a model without level was run, and all that was being asked was: *Is there a relationship between pre-test and post-test scores within each school?* Rather

² See, for instance, Hox, J. 2002, *Multilevel Analysis: Techniques and applications*, Lawrence Erlbaum Associates, Mahwah, N.J., pages 211–224.

than contrasting each school with a reference school, a different parameterisation of the model was run that allowed results to be returned for each school³.

Because the analysis began as a multilevel analysis within a Bayesian framework, it remained within the Bayesian framework (although this was not required) when it changed to a single-level analysis with 'school' as an independent variable. Markov chain Monte Carlo (MCMC) estimation with the Gibbs sampler was used. The analysis was run with a burn-in period of 5000 chains, with a monitoring period set to 50,000 chains. The priors were set to uninformative priors (uniform distributions for the fixed effects and inverse gamma distributions for the variance). The estimates presented in Table 6 are the modes of the posterior distributions, the standard errors are the standard deviations of the posterior distributions, and the 95% credible intervals are the boundaries of the central 95% of values of the posterior distributions. Modes are used in the same way as point estimates obtained from conventional maximum likelihood analyses, the standard deviations take the place of the standard errors, and the credible intervals take the place of confidence intervals.

The analysis returns a Bayesian Deviance Information Criterion (DIC), which can be used to compare models – a smaller DIC indicates a better fitting model. The DIC for the model without the nine slopes was 856.09, whereas the DIC for the model with the nine slopes fitted was 801.76. The difference (54.33) is large enough to indicate that the second model is a better fit to the data.

Figure A.1 shows the relationship between pre-test scores and post-test scores within each school. It shows the regression lines through the scatterplot of points (although the points are not shown in the figure) for each school. Table 6 (on page 18) also gives the slope of each school's regression line along with a 95% credible interval.

³ Different parameterisations of statistically equivalent models means that the same model is being applied. However, estimates of a different set of parameters are being sought.

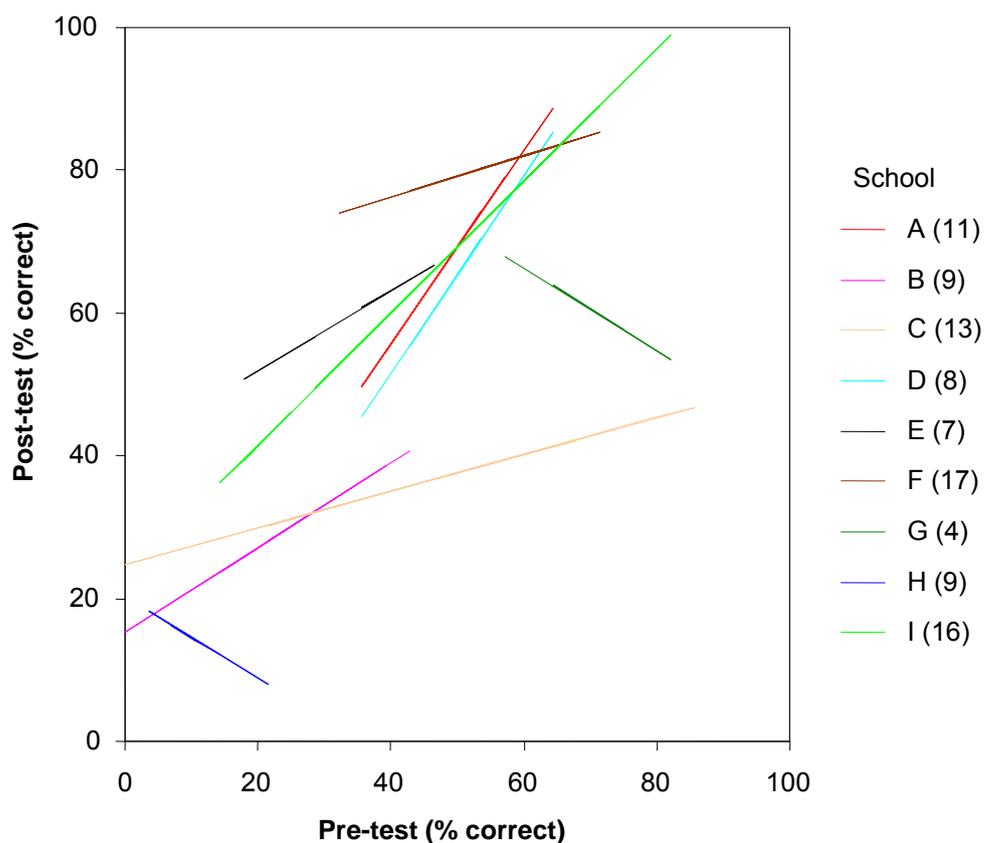


Figure A.1 Predicted regression lines for each school

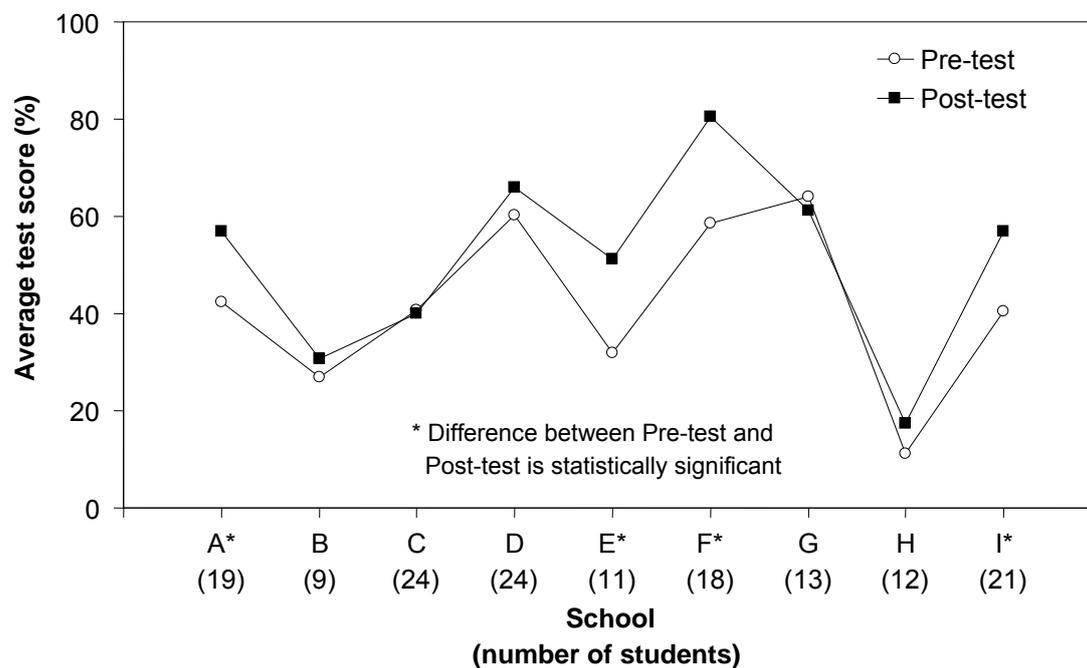
6.1.2 School level results: imputation methodology

The responses to three questionnaire questions, and pre-test and post-test scores were entered into SPSS's EM algorithm for imputation. Table A.1 and Figure A.2 show the same information as was shown in Table 5 and Figure 2 (both on page 15), except that Table A.1 and Figure A.2 show the information after imputation of missing post-test scores. There are some minor differences in the statistics and means reported, compared to Table 5 and Figure 2. However, the overall story remains the same for schools A, E, F and I – the difference between the mean pre-test and mean post-test scores are statistically significantly different.

Table A.1 The results of nine dependent groups t-tests (one for each school), testing for differences between pre-test and post-test scores (after imputation)

School	N	Pre-test	Post-test	t	df	p	Holms-adjusted α
A	19	42.3	56.9	5.73*	18	<0.001	.0063
B	9	27.0	30.6	<1	8		
C	24	40.8	39.9	<1	23		
D	24	60.3	65.9	2.10	23	0.047	.0100
E	11	31.9	51.1	3.90*	10	0.003	.0071
F	18	58.5	80.4	6.95*	17	<0.001	.0056
G	13	64.0	61.3	<1	12	0.565	
H	12	11.3	17.4	1.85	11	0.092	
I	21	40.5	56.8	3.44*	20	0.003	.0083

* t exceeds critical-t (ie, the difference is statistically significant)

**Figure A.2** Pre-test and post-test means for each school (after imputation)

Proceeding to the analysis where students' level is entered, as with the earlier analysis, some schools need to be dropped because they do not have any Level 2 students:

- School J has only one student

- Schools B, E and H do not have Level 2 students.

That leaves six schools in the analysis: A, C, D, F, G and I. For the students of these schools, there is:

- a significant effect for level (in each school, Level 2 students do better than Level 1 students)
- a significant effect for time in schools A, F and I (students do better on the post-test than the pre-test)
- a significant interaction, but only for School C.

Figure A.3 shows the pre-test and post-test mean for each level in the six schools after imputation for missing post-test scores. For all schools except School C the lines are, in effect, parallel (ie, there is no interaction between level and time). The result for School C remains counterintuitive. Note also that the imputation has resulted in some time effects not reaching significance. This result is also counterintuitive. It is expected that students would do better at the pre-test than at the post-test, even if Level 2 student had reached a ceiling at the pre-test.

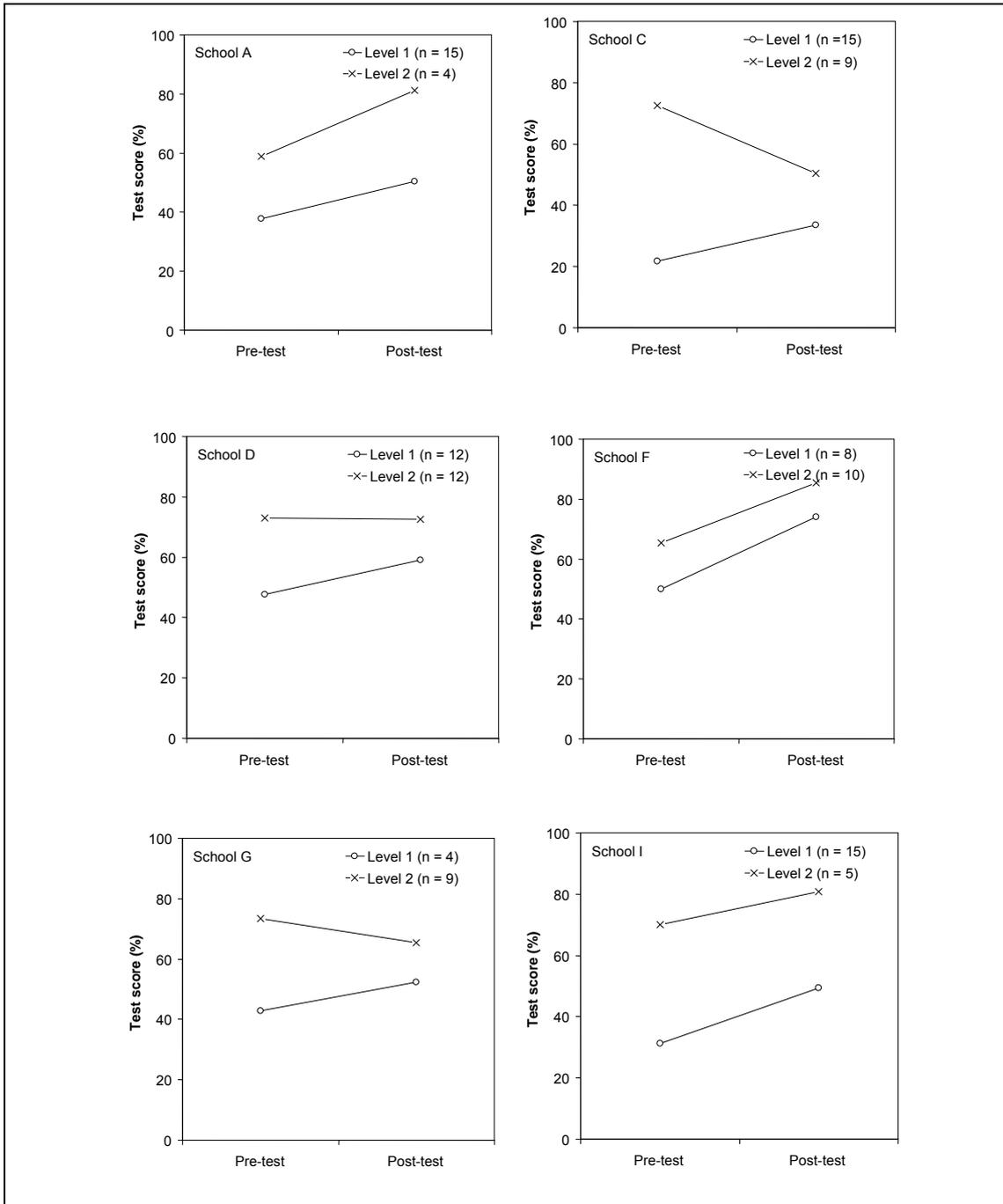


Figure A.3 Pre-test and post-test means for each level in each school (after imputation)

With respect to the MCMC analysis, there is not much change either. Table A.2 and Figure A.4 show the results after imputation (they correspond to Table 6 on page 18 and Figure A.1 on page 33). The difference between the DICs with and without the nine slopes fitted ($1349.69 - 1284.12 = 65.57$) is large enough to indicate that the model with the nine slopes fitted is a better fit to the data. Table A.2 shows that the 95% credible intervals (CIs) for five schools (A, C, D, E, I) do not span zero. Three of these schools (A, D and I) were the three from the earlier MCMC analysis with 95% CIs that did not span zero. The earlier MCMC analysis, suggested that the other two (Schools C and E), could have non-zero slopes.

However a lack of evidence prevented that claim from being made more forcefully. In the current analysis, there is more evidence, but the fact that much of this is imputed, not measured, prevents this claim from being made. In the earlier MCMC analysis, two schools (G and H) appeared to be going backwards, but it was argued that this appearance is likely to be the result of a lack of evidence. At least in the current analysis, the regression lines for these two schools are in the expected direction, but because the 95% CIs span zero, the relationship between pre-test and post-test scores in the two schools may well be zero. Note that for School G, the 95% CI only just spans zero, but in School H, the slope is small. The problem with School G, however, is that nearly all of its evidence is imputed.

Table A.2 MCMC estimates for the slope of each school's regression line through the scatterplot of pre-test against post-test (after imputation)

School	Slope	St Dev	95% CI
A (19)	1.12	0.26	(0.60, 1.63)
B (9)	0.61	0.44	(-0.25, 1.49)
C (24)	0.43	0.10	(0.22, 0.63)
D (24)	0.65	0.19	(0.28, 1.04)
E (11)	0.75	0.34	(0.08, 1.42)
F (18)	0.35	0.35	(-0.34, 1.04)
G (13)	0.38	0.25	(-0.11, 0.87)
H (12)	0.09	0.54	(-0.96, 1.12)
I (21)	0.95	0.15	(0.66, 1.23)

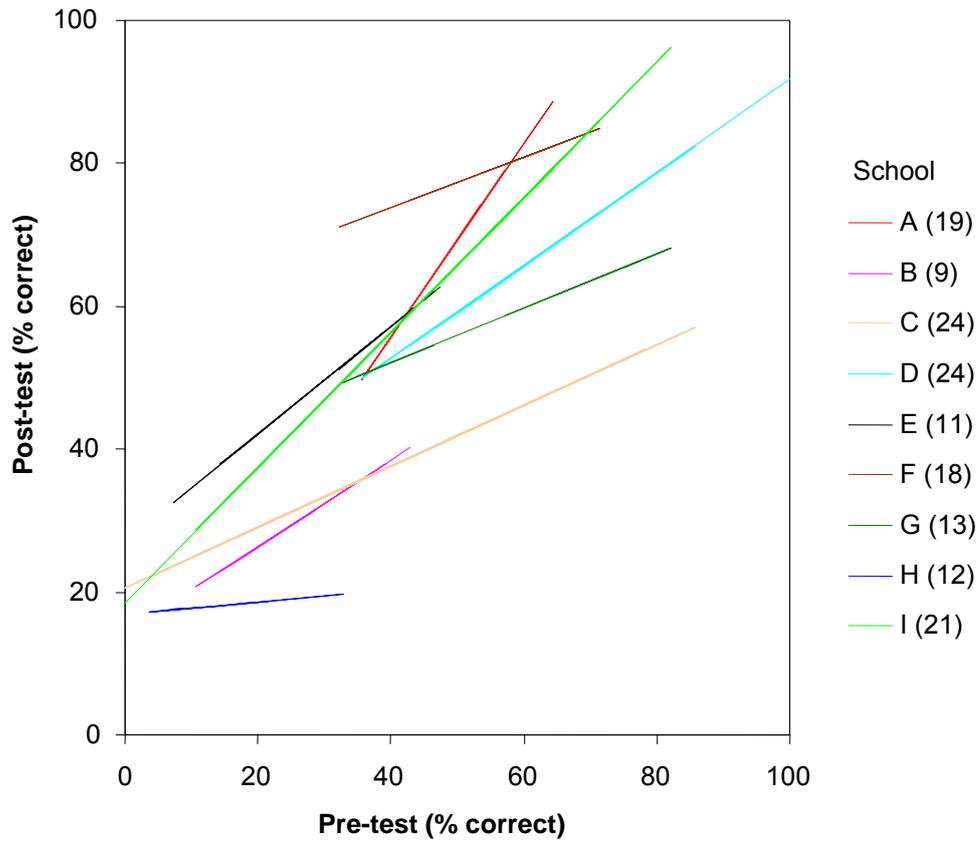


Figure A.4 Predicted regression lines for each school (after imputation)

6.2 Appendix 2: How it works

MOBILE APPLICATIONS FOR LANGUAGE LEARNING: How it works

The Le@rning Federation

Who	What	How
Students	Take classification test	<ul style="list-style-type: none"> Using the phones provided, call the system and follow the prompts to begin the test. Students will be prompted to enter the student number and then the PIN. Student number will be eight numbers and PIN will be four numbers. Each question should be answered as fully as possible. Students will be able to check their answers and re-record them if they wish before submitting them and continuing to the next question. Students will have no access to conversations until test is taken and marked.
Teachers	Grade the classification test	<ul style="list-style-type: none"> Log on to the website using teacher login and password and follow the links. Teacher username is the teacher name with no spaces and the password is seven characters. All conversations waiting to be marked will be listed in the unmarked conversations area. A headset/speakers will be required to listen to files from the site. Once mark and comment has been submitted it will appear on the student website.
Teachers	Enter result	<ul style="list-style-type: none"> Follow the marking guide to mark each question. Scores will be averaged. Score 1–60 will allocate student to Level 1 Score greater than 60 to allocate student to Level 2 Press ‘Save marks’ button. Result will appear in student area.
Teachers	Advise students	<ul style="list-style-type: none"> All conversations will be available to students once classification has been activated. Teachers may wish to direct topic order with students.
Students	Start conversations	<ul style="list-style-type: none"> Using the phones provided, call the system and follow the prompts to begin the test. Students will be prompted to enter the student number and then the PIN.

		<ul style="list-style-type: none"> • Each question should be answered as fully as possible. • Students will be able to check their answers and re-record them if they wish before submitting them and continuing to the next question. • Students may do more than one conversation at each session.
Teachers	Review student conversations	<ul style="list-style-type: none"> • Log on to the website using teacher username and password and follow the links. • All conversations waiting to be marked will be highlighted. • Review and comment on work. • Once reviewed conversation has been submitted it will appear on the student website.
Student	View teacher review	<ul style="list-style-type: none"> • Log on to the website using student login and PIN and follow the links. • Recently reviewed work will be highlighted. • Model answers will be available for questions already answered.
All	Forum site	<ul style="list-style-type: none"> • Access the forum site to see what other students/teachers are saying about the program. • Add site to favourites in browser and visit often.

6.3 Appendix 3: Questionnaire instrument content and participant group

Key research question: Is the Mobile Application for Language Learning approach a viable option for teaching and learning of languages?					
<i>Aspect</i>	<i>Teacher questionnaire</i>	<i>Student questionnaire</i>	<i>School leadership questionnaire</i>	<i>Teacher semi-structured Interview</i>	<i>Student semi-structured Interview</i>
How does the level of available connectivity (to web-based services) impact on this type of application?	√		√	√	
How effective would a model such as MALL be in supporting distance education teachers in implementing a language course?	√				
How effective would a model such as MALL be in supporting teachers who work across a number of schools in implementing a language course?	√				
Which aspects of the MALL can you envisage transferring well to other language learning areas?	√				
Was the MALL system and mobile phone 'lock down' successful in addressing issues of security and privacy of participants?	√		√		
What are the perceived issues inherent in using mobile	√		√		

technologies in terms of security, privacy and child protection perspectives?					
What aspects of using a model such as MALL would require additional professional learning opportunities to support its effective implementation?				√	
How often do you use a mobile phone in your everyday life outside of school?		√			
How often do you use a mobile phone at school for schoolwork?		√			
If you have your own mobile phone how do you feel about having to carry two phones?		√			
How much of an issue is access to internet and computers for the success of such a program?	√		√		
What general issues may get in the way of implementing a program such as MALL at your school?	√		√		
Does your school currently have a policy on ICT use?			√		
What are the key aspects that should be included in a school policy on ICT use?			√		

Does your school currently have a policy specific to the use of mobile phones and other mobile technologies?			√		
What aspects specific to the use of mobile phones and other mobile technologies should be included in a school policy on ICT use?			√		
What are the key considerations that need to be taken into account when formulating a whole-school policy/approach to the implementation of mobile technologies?			√		
How frequently did you use the MALL online forum site?	√				
Did you find the forum a useful way to communicate to others involved in the project?	√				
Did you post any questions on the forum?	√				
Did you post any responses to questions on the forum?	√				
Do you generally participate in other educational forums or networking sites?	√				
Would you use a similar forum site if it was available for other subject areas	√				

(ie non-language learning focus)?					
How could the forum site be improved?	√			√	
Have you used The Le@rning Federation digital curriculum resources previously?	√	√			
How much of a problem was it for you being responsible for the mobile phone during the project?		√			
Where did you access the internet for this project: at school, at home, somewhere else?		√			

Key research question: Is this approach a viable option for inclusion in the overall blend of teaching languages to students?					
Aspect	Teacher questionnaire	Student questionnaire	School leadership questionnaire	Teacher semi-structured interview	Student semi-structured interview
Have you noticed improvements in student engagement for the MALL model over other teaching methods?	√			√	
Have you noticed any improvements in student 'time on task' for the MALL model over other teaching methods?	√			√	
Have you noticed any improvements in terms of improved learning outcomes for the MALL model over other teaching methods?	√			√	
Would it be desirable to integrate MALL with other teaching resources?	√				
How could the MALL be best integrated with other teaching resources such as Asian Education Foundation resources?	√				
How could MALL be best integrated with other technologies (eg podcasts)?	√			√	
Does the MALL model support collaborative learning	√			√	

opportunities with Asian counterparts?					
Looking to the future, what might the next phase in implementation of a system such as MALL be across a sector?			√		
How did you feel about speaking to the computer system and answering the questions?		√			√

Key research question: Has this approach improved the language skills of the participating students?					
Aspect	Teacher questionnaire	Student questionnaire	School leadership questionnaire	Teacher semi-structured interview	Student semi-structured interview
How long have you been learning Indonesian?		√			
How much did the computer system help you to improve your Indonesian overall?		√			
By how much did using this system help you to improve your Indonesian listening skills?		√			
By how much did using this system help you to improve your Indonesian speaking skills?		√			
By how much did your confidence level in being able to respond to questions in Indonesian improve?		√			√
How would you define 'success' for the MALL project at the individual student level, the whole class, small group and overall?	√			√	
According to your criteria for success, rate the success of this program at the: individual student	√				

level, the whole class, small group and overall.					
Which aspects of the MALL would you rate most highly/least favourably in achieving the goal of students learning Indonesian language?	√			√	