Comparison of English versus Hindi medium students for programming abilities acquired through video-based instruction

Yogendra Pal
Department of Educational Technology
Indian Institute of Technology Bombay
Mumbai, India
e-mail: yogendra.pal3@gmail.com

Sridhar Iyer
Department of Computer Science and Engg
Indian Institute of Technology Bombay
Mumbai, India
e-mail: sri@iitb.ac.in

***Abstract*— Acquisition of programming skills is a key aspect of computer science curricula. Students who have studied in their local language have significant difficulty in acquiring programming skills through English medium instruction. Given the known benefits of local language instruction, one approach to tackle this problem is to record the lectures and provide translated videos to such students. As yet, there is no experimental data on the effect of teaching programming using translated videos, on student *achievement*. In this paper, we report our findings from a study of three groups of students learning programming by watching 3-hours of video lectures on introductory programming. One group of Hindi medium students watched the videos in English while a similar group watched the corresponding Hindi videos. As a baseline, one group of English medium students watched the English videos. We found that the difference between the three groups on post-test achievement was not significant.**

Keywords- Programming ability; local language learners; Scaffolding; video instruction.

I. Introduction

Acquisition of programming skills is a key aspect of computer science curricula. The ACM/IEEE-CS joint task force has allotted 20 hours to programming, out of a total of 142 hours for the entire undergraduate computer science curriculum [15].

India has a large number of students who study in their local language upto 12th standard and then have to adapt to English language instruction for their undergraduate education, since most of the higher education institutes use English as the medium of instruction [1]. These students have significant difficulty in acquiring programming skills, as evidenced by their low success in the university exams [2,11].

To tackle this problem, one approach adopted by the National Mission on Education through ICT [3] is to record video lectures delivered by experts, translate the videos into local languages, and disseminate the videos [4,5]. Some benefits of using video lectures, such as low cost and high availability, are well-known [19, 20]. However, there are no significant studies to determine the effectiveness of video tutorials for student achievement.

Also, there is a lot of literature on the benefits of local language instruction [6]. There is also work on the effectiveness of using local languages for classroom teaching of computer science subjects [7,8,9] and programming [10]. However there is no experimental data on the effect of teaching programming using translated video tutorials, on student *achievement.* Towards this end, we conducted the following study.

We selected three 1-hour video lectures on introductory programming, delivered in English by a well-known professor. The professor repeated the lectures in Hindi to create the translated videos. We then chose a reputed engineering college that had a mix of students coming from English and Hindi medium schools. We created 3 groups of 31 students each. The control group had Hindi medium students watching the three lecture videos in English. The experimental group had Hindi medium students watching the corresponding Hindi videos. As a baseline, we had English medium students also watching the English videos. For our choice of topics and treatment, we found that the difference between the three groups on post-test achievement was not significant.

In Section II we present the related work on teaching programming through local language instruction. In Sections III and IV, we give the details of our research questions and methodology, respectively. The results are in Section V, followed by discussion in Section VI.

II. Related Work

Experiments for teaching using a combination of English and local language have been carried out for subjects like mathematics, physics and chemistry [13,14]. These studies show that there is a benefit in using local language for instruction. In [13], the authors conducted a comparative study on English and Korean medium students and found that Korean medium students have greater achievement in mathematics if taught in Korean. In [14], the authors conducted an experiment on teaching physics to two groups of students and found that using the local language (Tamil) results in a significant increase in test-score after treatment.

Programming is taught in various ways, such as classroom lectures along with laboratory assignments [22], computer aided learning (CAL) [17], online learning [16] and blended learning environments [18]. We chose to explore the use video lectures for teaching programming.

978-0-7695-4759-6/12 $26.00 © 2012 IEEE
DOI 10.1109/T4E.2012.30

Some benefits of using video lectures, such as low cost and high availability, are well-known [19, 20]. Students and teachers also agree that such videos should be translated in local languages [21].

However there is not much work on the effectiveness of translated video lectures for student achievement, nor on teaching of programming in local languages. We found only one work [10], that suggests a bilingual model for teaching programming to undergraduate students in China, but it is still being implemented and no experimental data is available. Moreover it uses a classroom-based model, unlike our video-based approach. So we feel that our study would be a useful contribution.

III. Research Questions

We use the term "medium" to denote the medium of instruction in 12 years of schooling. In our experiment, the medium could be the same as the local language (Hindi) or different (English). We use the term "MoI" to denote the medium of instruction in the treatment. In our study, the MoI for the video lectures is either English or Hindi.

At a broad level, our question is: What is the impact of the MoI on the programming abilities of learners from various mediums? This is operationalized into the following specific questions:

1) Do undergraduate Hindi medium students learning introductory programming by watching three 1-hour video lectures in Hindi, perform better than similar students who watch the same lectures in English?

2) Does self-reported prior knowledge play a role in the students' performance?

IV. Methodology

A. Sample:

The sample consisted of 93 engineering 1st year undergraduate students of a reputed college in North India. Convenience sampling was used as joining the training was voluntary.

The sample was divided into 3 groups according to their prior medium, as shown in Table 1.

TABLE I. Medium of Instruction for various groups

Medium in 12th Std	Treatment	Group (Name)	N
English	English	Baseline (EE)	31
Hindi	English	Control (HE)	31
Hindi	Hindi	Experimental (HH)	31

B. Data collection:

To measure programming ability, performance scores on a post-test were collected. To determine prior academic achievement levels, overall percentage of marks in 12th Std were collected.

C. Instrument:

We used a 3-item survey to collect data about students' background. The items for each student were: (i) MoI in 12th standard (English or Hindi), (ii) Overall percentage of marks in 12th standard, and (iii) Whether they have prior knowledge of programming (yes or no).

We used a paper-based post-test containing 14 multiple choice questions, each having exactly one correct answer. We looked for a concept inventory for programming but found that the standardization of assessment instruments for programming ability is still ongoing [12]. So we created the post-test based on questions that typically appear in the University exams and those given in standard textbooks. We included only those questions that directly mapped to the learning objectives in our video lectures. We had 14 questions, of which 8 were at the Knowledge and Comprehension level and 6 were at the Apply and Analysis level [23]. One sample post-test question at each level is shown below :-

(Knowledge level) Q1. In the statement int i='a';
What will be stored as value of variable 'i'?
a) ASCII value of character 'a' [ ] b) 'a' itself [ ]
c) Program will return an error [ ] d) None of them [ ]

(Apply level) Q6. Find the value of i in:
int i = 20 + 6 / 4
a) 21.5 [ ] b) 21 [ ] c) 6.5 [ ] d) None of them [ ]

D. Procedure

1) Survey: We first conducted the survey and then divided the students into 3 groups, based on the MoI of their 12th Std, as shown in Table 1. We compared the means of the 12th Std marks for the groups (EE = 63.73, HE = 61.67, HH = 63.44) and found them to be equivalent. We did ANOVA to confirm the equivalence. There was no pre-test.

2) Arrangement: We arranged separate classrooms for the 3 groups. Each classroom was equipped with projector and sound system.

3) Treatment: We used only recorded video lectures for the treatment. Our rationale was to eliminate any bias that may occur due face-to-face communication, such as the instructor adapting the lecture dynamically based on cues from the students. The videos were already publicly available, created under the National Mission of Education through ICT [3]. They were created by recording the classroom lectures given for the CS 101 course at IIT Bombay. The videos not only contain the lectures and slides, but also the screen-capture of demonstrations for writing, running and compiling computer programs. The English recordings were created during normal class hours while the Hindi recordings were created during extra classes specifically for Hindi medium students. Although the lectures in the two languages are not completely identical translations, they addressed the same Learning Objectives (LOs) in nearly the same manner. The list of LOs for the chosen lectures is given below in Table II.

Each group watched three 1-hour video lectures, with breaks after each video, over a total period of five hours. There were no additional tutorials or laboratory exercises. The MoI in the videos was Hindi for the experimental group and English for the other two groups (See Table 1). The lectures were on the topics: (1) Introduction to Computers, (2) Introduction to Cpp, and (3) Numerical Computing.

TABLE II. LEARNING OBJECTIVES

Sr. no.	LO (Students should be able to...)	Video no.	Q. no.
1	Describe the need and importance of computer programming.	1	-
2	Identify all valid/ invalid characters in a given cpp program.	1	-
3	Identify valid/ invalid arithmetic operators.	1	14
4	Identify how much memory is reserved for a variable of a given data type.	2	2
5	Differentiate between initialization and declaration statements in a given program.	2	4, 5
6	Identify valid/ invalid use of variable names, keywords in a given cpp program.	2	14
7	Identify valid/ invalid use of brackets in arithmetic instructions.	2	12
8	Calculate an expression according to the given data types.	2	6
9	Write a mathematical equation using appropriate computer programming expression syntax.	2	7
10	Distinguish between valid/ invalid programming expressions.	2	14, 11
11	Solve a programming expression by executing each operator according to it's order of precendence.	2	6, 13
12	Determine storage for character variables.	3	8
13	Identify valid/ invalid constants for basic data types.	3	10
14	Describe what will be stored in memory for a given type of variable.	3	1, 9
15	Choose appropriate data type for a variable given it's use in a programming problem.	3	3

4) Posttest: To investigate the effect of the MoI on achievement scores, we conducted a post-test using the instrument we had designed earlier. The mapping of post-test questions to learning objectives in the videos, is also shown in Table II. Each student had to attempt the post-test individually, within a time limit on thirty minutes. There was no negative marking.

5) Analysis: We performed quantitative analysis of the post-test scores for the different groups. We computed the means for each group. We used ANOVA to determine which means are significantly different from one another. We used factorial analysis to determine the effect of prior knowledge on the post-test scores.

V. RESULTS ANALYSIS

The distribution of post-test scores for the three groups (EE, HE and HH), and their means are shown below.

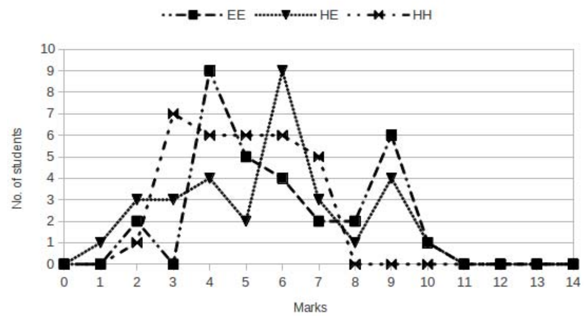

Fig I. MARKS DISTRIBUTION OF STUDENTS IN EACH GROUP

TABLE III. MEANS OF STUDENTS IN EACH GROUP

Group	N	Mean (out of 14)	Percent -age	Std. dev.	Std. error Mean
HH	31	4.77	34.07 %	1.50	0.27
HE	31	5.52	39.43 %	2.37	0.43
EE	31	5.90	42.14 %	2.25	0.41

As seen from Fig I, the representation of high scores (greater than 7 out of 14) is low in all the three groups, particularly HH. From Table III, the percentages of the means are low and the difference in means does not appear to show a difference between the groups.

We performed ANOVA to determine the variation between groups but did not find the differences to be significant, as shown in Table IV.

TABLE IV. ANOVA FOR THE 3 GROUPS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	20.409	2	10.204	2.356	.101
Within Groups	389.871	90	4.332		
Total	410.280	92			

To investigate at a finer level, we analyzed the performance of the groups in two categories of questions. Category 1 consisted of questions at the Knowledge or Comprehension level (Q1, Q3, Q6, Q7, Q8, Q11, Q13), while Category 2 consisted of questions at the Apply or Analysis level (Q3, Q6, Q7, Q11, Q13, Q14). The number of students who answered the questions correctly in each group is shown below (Fig II). We did not observe any clear pattern in the Category 1 questions. In Category 2 questions we saw that for each question number of students who answered them correctly is highest in group EE, followed by HE and HH.

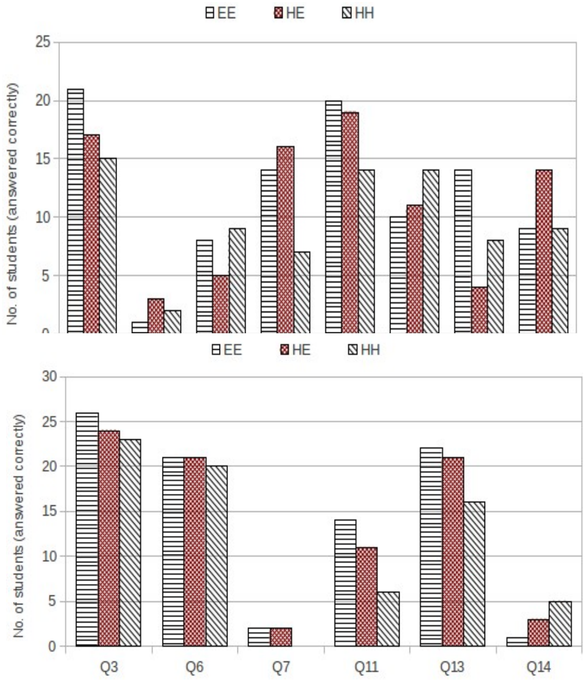

Fig II. Performance in the two categories of questions

In order to see if there is a difference between the post-test score of sub-groups that have prior-knowledge (PK) of programming, we determined the means for various sub-groups as given in Table V.

TABLE V. Sub-group means

Students	Stats.	EE	HE	HH
Total Students	N M SD	31 5.90 2.25	31 5.52 2.37	31 4.77 1.50
Has prior knowledge of programming (PK)	N M SD	16 6.63 2.09	12 5.42 3.43	5 5.20 1.17
No prior knowledge of programming (NPK)	N M SD	15 5.13 2.09	19 5.58 2.52	26 4.69 1.48

Although the prior-knowledge is self-reported and not determined through a pretest, we found that it plays a role when the medium and MoI are matched (EE and HH groups), with mean scores for EE being greater than HH. To delve further, we carried out pair-wise paired sample t-test between the sub-groups within each group. We found there is no difference between the sub-groups for any group at 0.05 significance level.

To investigate further, we plotted the interaction effect between MoI and prior knowledge for Hindi medium (HE and HH groups), which is shown in Fig III.

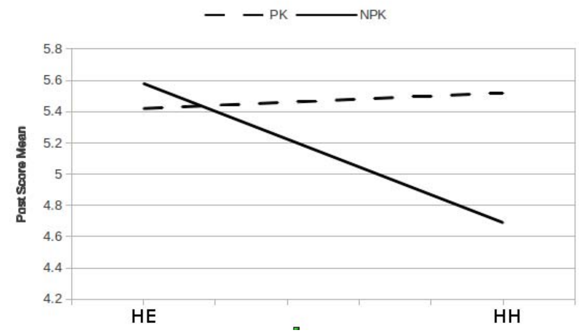

Fig III. Interaction plot of Prior-knowledge with MoI for Hindi medium

Since there appears to be interaction between the factors, we carried out a factorial analysis of variance using ANOVA but did not find them to be significant (Table VI).

TABLE VI. Factorial Analysis of Variance using ANOVA

Source	df	SSQ	MS	F	p
MoI	1	3.25	3.25	0.80	0.38
PK	1	0.32	0.32	0.08	0.78
MoI x PK	1	1.20	1.20	0.29	0.59
Error	58	235.89	4.07		
Total	61	240.66			

VI. Discussion And Conclusion

Our treatment of teaching introductory programming using videos of 1-hour lectures in English versus Hindi, did not result in any significant difference between post-test scores of experimental, control and baseline groups. One reason for the lack of difference could be that the mean scores are consistently low for all three groups.

Some reasons for the mean scores being low could be:

1) *Treatment:* It is difficult for students to be actively engaged while watching three consecutive 1-hour lecture videos, even with breaks. Moreover, all the students had to watch the video simultaneously, without any scope for watching at their own pace. Hence the actual learning for any group, as measured by post-test scores, was low.

2) *Instrument:* The lack of standardized assessment instruments is a major hurdle. Although our post-test was based on questions from the University exams and the standard textbook, it is possible that they were difficult for the students, either at a conceptual level or with respect to the time limit (2 minutes per question).

3) *Topic:* The topics chosen were of introductory nature, so it is possible that the corresponding learning objectives and questions are not much influenced by MoI.

4) *Subject:* Computer programming is inherently an application-oriented skill and needs laboratory / tutorials. So a treatment of only watching lecture videos may not be appropriate for this topic, leading to low scores for all the groups.

Hence, generalizing our results to conclude that providing translated videos to local language learners is ineffective, is also not correct. More experiments are required in order to determine: (i) the type of topics as well as the duration of the videos for which such treatment could be effective, and (ii) type of treatment for programming and duration that would be effective for local language learners. In our future work, we propose to address the latter, by various experiments.

References

[1] SK Mitra (2010). "Internationalization of Education in India: Emerging Trends and Strategies." In the Journal of Asian Social Science, June 2010, pp. 105-110.

[2] Prabhakar Kaushik and Dinesh Khanduja (2010). "Utilising six sigma for improving pass percentage of students: A technical institute case study." In the journal of Educational Research and Review, September 2010, pp. 471-483.

[3] http://www.sakshat.ac.in/ as seen on 04. mar. 2012

[4] http://nptel.iitm.ac.in/ as seen on 04. mar. 2012

[5] http://spoken-tutorial.org/ as seen on 04. mar. 2012

[6] Michele S. Moses (2001). Why Bilingual Education Policy is Needed: A Philosophical Response to the Critics. Bilingual Research Journal, 24(4), 333-354.

[7] Leo Gomez, David Freeman and Yvonne Freeman. Dual Language Education: A promising 50-50 model. Bilingual Research Journal, volume 29, Issue 1, 2005. pages 145-164

[8] Jinshu Han (2010). Bilingual Teaching Practice of Undergraduate Computer Graphics. Second International Workshop on Educational Technology and Computer Science. Pages 441-444.

[9] Shuang Liu, Xizuo Li and Li Zuo (2011). Adopting communicative teaching method in computer major bilingual teaching. Theory and Practice in language studies, vol 1, no 2, pp. 187-190, February 2011

[10] Hanjing Li, Kuanquan Wang, Yuying Wang. A bilingual teaching modal in a programming language course. Second international conference on education technology and training, 2009, pages 197-200.

[11] Jens Bennedsen and Michael E. Caspersen (2007). "Failure Rates in Introductory Programming." In SIGCSE bulletin, June 2007, pp. 32-36.

[12] Allison Elliott Tew and Mark Guzdial (2011). The FCS1: A Language Independent Assessment of CS1 Knowledge." SIGCSE, Mar. 2011, pp. 111-116.

[13] Dong-Joong Kim, Joan Ferrini-Mundy, Anna Sfard (2012), How does language impact the learning of mathematics? Comparison of English and Korean speaking university students' discourses on infinity", International Journal of Educational Research. Feb. 2012.

[14] D. Venkatesan, Dr. RM Chandrasekaran and Dr. A. Velangani Joseph (2010), "Significance of Medium of Instruction and Personality in Designing Adaptive e-Learning Systems.", In the International Journal of Computer Applications, Oct. 2010, pp. 25-28.

[15] Computer Science Curricula 2013 Strawman Draft (February 2012), page 33

[16] http://www.w3schools.com/ as seen on 02. Mar. 2010

[17] Ahmad S. Mashhour (2012), "A Knowledge-Based Educational Module for Object-Oriented Programming & The Efficacy of Web Based e-learning.", In the International Journal of Advanced Computer Science and Applications, Feb. 2012, pp. 157-165.

[18] Said Hadjerrouit (2008), "Towards a Blended Learning Model for Teaching and Learning Computer Programming : A Case Study", In Informatics in Education, 2008 Vol. 7, No. 2, pp. 181-210.

[19] Kathleen Barnes, John Bowers, Julie H. Dent (2009), "Video Tutorials: A Sustainable Method for Campus Technology Training", In EDUCASE Quarterly (EQ), 2009 Vol. 32, No. 3.

[20] http://117.211.119.206/jspui/handle/123456789/2357

[21] J. Ravi, H.J. Jani (2011), "A Critical Study of NPTEL", In International conference on Technology for Education (T4E), Aug. 2011, pp. 35-42

[22] http://ocw.mit.edu/index.htm as seen on 05 mar. 2012.

[23] David R. Krathwohl "A Revision of Bloom's Taxonomy: An Overview", In Theory into Practice, June 2010, vol. 41, No 4. pp.212-218