

Chapter 5

Conclusion, Discussion, and Recommendations

5.1 Summary of Procedure and Findings

The first objective of the study was to determine whether frequent assessment of upper secondary second language learners would impact final ability differently based upon the relative proficiency, age and gender of learners. The second objective of the study was to determine whether the same learners would form different opinions regarding the course and its component parts, including instruction, the instructor, assessment and their own confidence. The population was limited to upper secondary learners of English in northern Thailand with limited proficiency, ages 13–19, but mostly 14–17 and 416 upper secondary Thai students participated in the study. Learners were randomly assigned by class section to treatment and control groups to form the basis for comparison. Treated learners received short, predominantly multiple-choice review treatment instruments at the end of every lesson while untreated learners went without assessment. Instruction otherwise remained uniform for all learners. Prior to the instructional phase *The English Placement Examination* was issued to all learners to determine relative proficiencies prior to instruction. During the instructional phase, treated learners received a series of three *Treatment Instruments*. The relative English language ability of learners was measured using an end of term final assessment of ability, *The English Language Placement Examination*, and results were compared based upon whether learners had or had not received assessment or not as well as in reference to gender, age and proficiency. How learners responded to individual sections of the test was also analyzed. Opinions toward English language learning were measured using the *Opinion Survey Questionnaire*. A *Field Journal* was kept to record incidental observations. Data was analyzed relative to a purposeful sample taken from the total pool of participants with 25 learners representing each of four ranges of ability for both treated and untreated groups, yielding two proficiency-neutral groups of 100 participants each, the first representing treated learners, the second, the control. Proficiency corresponded to measurements taken on the pre-instructional placement examination. Results revealed no statistical differences between treated and untreated learners relative to the entire data set. Highly proficient learners scored significantly higher when assessments were withheld from lessons. Younger learners scored significantly better when assessments were included in instruction. Older learners scored significantly better when assessments were withheld from instruction. Female learners scored significantly better on all measurements, but not as a result of the treatment. Opinions toward English language learning did not vary based upon frequency of assessment given.

5.2 Discussion

Motivating Questions

Returning to the broad, theoretical questions that informed and inspired the design of the study, whether assessment can interfere with language learning and for whom seems to have been demonstrated by this research: older learners who received assessments performed significantly less well than learners of the same age who did not receive assessments and the same was true of learners who began the course with a relatively high proficiency. Whether the assessments were too easy and caused boredom for learners or whether any similar variety of assessment would create the same drag on older and more proficient learners is not clear and would require further, more in-depth research. That the assessments created a barrier to learning seems evident. Whether assessment necessarily improves language learning was at least partially answered: assessment does not necessarily improve language learning—in certain cases, as with the older learners in this study, it appears to have interfered with learning. Whether assessment remains essential to language learning was not and likely could not have been answered by this study. Nevertheless, the assessment strategy used provided some advantage for younger learners and a marginal, though statistically not significant advantage for lower proficiency learners. Female learners tended to score higher in general but the data may suggest that female learners were better off without the review assessments—perhaps owing to their higher relative proficiency—and that males may have benefitted more from inclusion of the assessments, discussed in greater detail below.

Proficiency

Untreated learners in the highest proficiency range scored significantly higher than treated learners in the highest proficiency band, a difference that was even greater amongst female learners in the highest proficiency band. This could possibly suggest that the addition of assessments had the reverse of expected effect on the highest proficiency learners and possibly served as anchors to their growth. Female learners in the highest-proficiency group may have been especially disadvantaged by the presence of assessments and the absence of other activities their presence required. This corresponds with research that suggests that advanced learners may grow dispirited by classes that are too easy (Moon, Brighton, and Callahan, 2003). Treated learners in the lower three bands of proficiency scored consistently but not significantly higher than untreated learners in the same proficiency range.

Age

Younger learners appeared to have benefitted from the treatment. Not only did treated younger learners score significantly higher than untreated younger learners, they scored higher than older treated learners as well, though not significantly; on nearly every other measurement, older learners had scored significantly higher. The treatment disadvantaged older treated learners who scored less well than their untreated counterparts, similar to the disadvantage that was revealed for high proficiency learners. It is possible and the means suggest that the treatment assessments were adequately designed for younger learners. Research has shown that

the wrong type of classroom instruction, represented here by the potentially inappropriate assessment strategy, can in fact impede language learning (Krashen and Seliger, 1975 and Long, 1983 in Bley-Vroman, 1990).

With respect to age and acquisition, research has generally shown that younger teenagers, ages 12 to 15, can achieve second language competency at a greater initial rate than learners just older and younger (Bley-Vroman, 1990). Younger learners in this study, who were between 13 and 15, fell into this optimal range and may have simply benefitted more from instruction overall. The control framework may have contained a dearth of input in lieu of the inclusion of assessment. Older learners (16–19) fell outside of this special category and may have therefore suffered from a generally less successful acquisition experience. However, older learners scored higher than younger learners prior to any instruction and younger learners who received the treatment scored better than younger learners who did not, suggesting the treatment was at least partially responsible for a more pronounced relative ability. Other research has suggested that some younger learners may respond poorly to assessment when given early (Shaaban, 2001), theoretical underpinnings that have been operationalized in Finland and Norway with good results (Hasselgren, 2000; Moore, 2013), where assessments are not given to younger learners, but may refer to learners younger than those included within this study.

Gender

In general, female learners scored significantly higher. Some of the data collected here suggests that female learners performed better without the assessments used and that male learners performed better when assessments were used, but the specific comparisons reveal no statistically significant differences between treated male and female learners. These findings fall in step with general findings in language education gender studies, which predict variances based upon gender (Oxford, 1993) and also with the observation that female English language learners in Thailand tend to be more highly motivated on average than male learners (Khamkhien, 2010).

Opinions

Differences in opinions towards English language learning between learners who were assessed and learners who were not assessed were not remarkable or statistically significant, and neither were they remarkable or statistically significant when factored by age or gender. This suggests, at the very least, that the assessments did not affect the opinions learners formed about the course. It could also reflect that rules of politeness in cultures like Thailand make direct criticism of a teacher improbable, which has been reflected in some studies (Belchamber, 2013).

Researcher Observations

The assessment strategy served a useful purpose in terms of organizing younger classes. Younger learners were perhaps accustomed the participating in assessment-based activities and consequently were able to organize themselves more efficiently when assessments were included as a feature of pedagogy. Older learners manifested explicit displeasure with the assessments in one class section and otherwise tended

not to respond enthusiastically to assessments. Certain younger learners were eager to learn their scores on assessments whereas older learners were not particularly interested in their performances.

5.3 Pedagogical Implications

Assessments that are not specifically calibrated for a learning group can impede rather than promote acquisition. Instructors should only include assessments and assessment activities that have been specifically calibrated to the learning set, as recommended in the majority second language assessment literature (Bachman and Palmer, 1996) as well as providing activities that are flexible in nature and can accommodate varying learner-types (Laing and Kamhi, 2003). The results suggest that younger learners may benefit more from structured activities whereas older learners may thrive in discussion-based classes, and the distinction here was made between ages 13–16 and 16–19. Learners sixteen and above, at least within this model, seemed to benefit from a looser structure and may therefore benefit from classes that feature greater exposure to materials and discussion. Such information may support the use of needs and situational analyses (Richards, 2001; Graves, 1999).

5.4 Limitations of the Study

The study was limited predominantly by the scarcity of time and the availability of supporting resources. An ideal study of the same nature would measure the variances in student performance across a longer period of time, for example an entire semester of study, to determine whether various learner-types performed differently depending upon whether assessment was included within the course of instruction or not. In order to measure the long-term effect of the treatment, a second posttest and a secondary analysis of opinions would have been useful. The study only measured a single form of “traditional” assessment, multiple-choice treatment instruments with some short answer items, as compared to the absence of assessment, and did not take into account the effect other forms of assessment might have had as compared to the absence of assessment or, perhaps more interestingly, an alternative form of assessment or a different pedagogical structure. Focus group interviews and more thorough data gathering relative to the personal characteristics and affective factors of participants would have provided useful data on opinions, reactions and possibly reasons for variances in performance.

5.5 Recommendations for Future Research

Subsequent research would do well to consider relationships between assessment, age, gender and proficiency in greater depth with respect to the relative effectiveness several varieties of assessment yield on wider varieties of learners and learner-types to more effectively isolate particular modes and frequencies of assessment provide benefit to language acquisition. Future research should analyze the differing effects of varying degrees of consequence on learner performance and the quality and rapidity of acquisition. More concerted and in-depth analysis of learner attitudes and feelings toward assessment as well as the development of more purposeful and detailed learner

profiles would provide more interesting results. Researchers might also take more careful consideration of learner motivation between groups.

5.6 Conclusion

Different learners respond differently to different learning conditions and tasks. That many learners in the broader fields of education feel claustrophobic in the modern educational paradigm, which is defined by confinement and external management, is not surprising. Of the learners measured for this study, those who had originally scored in the highest category of proficiency fared better on the final examination when intermediary assessments were withheld. Younger learners seemed to benefit somewhat from the assessments whereas older learners may have been disadvantaged. Opinions were not affected by the assessment strategy. It seems clear that older and more proficient learners would benefit more from more complex tasks.

PAYAP UNIVERSITY