The Relationship between Self-Efficacy and Language Learners’ Grades

Stephen A. Templin
Meio University

This research explores the hypothesis that students with high self-efficacy: high beliefs in their capabilities to accomplish a task, will achieve higher grades in second language classes than students with low self-efficacy. Seventy-four Japanese high school students were asked to fill out a questionnaire and indicate by a yes or no response which grades they thought they could attain. They also rated their degree of confidence as a percentage for each level. Participants’ scores were the total of confidence percentages for “yes” answers. In estimating reliability, Cronbach’s alpha for the questionnaire and its subsections was .96, .98, and .91 respectively. A t-test was used to determine if there was any significant difference between low and high self-efficacy students’ grades. High self-efficacy students achieved significantly higher grades than low self-efficacy students.

Self-efficacy is belief in how well one can accomplish tasks. Although self-efficacy studies have appeared frequently in psychology (Bandura, 1986; Lee & Bobko, 1994; Locke & Latham, 1990) and management research (Gist & Mitchell, 1992; Gist, Schwoerer & Rosen, 1989; Matsui, Ikeda & Ohnishi, 1989; Matsui & Tsukamoto, 1991), self-efficacy research in second language acquisition (SLA) is rare.

Self-efficacy is important because it influences an individual’s performance in two ways. First, a person with high self-efficacy towards a
task pays more attention, makes a greater effort, is more persistent, and uses a greater variety of strategies to accomplish a task than one with low self-efficacy (Earley & Lituchi, 1991; Lee & Bobko, 1994). High self-efficacy individuals attribute failure to internal causes more than low self-efficacy individuals, who prefer to blame external events (Earley & Lituchi, 1991; Lee & Bobko, 1994). Consequently, when those with high self-efficacy encounter obstacles, setbacks, and failure, they will increase their attention, effort, persistence, and strategies in order to accomplish the task. In contrast, those with low self-efficacy are more likely to give up when faced with similar obstacles.

Second, highly efficacious people actively seek challenging goals and these goals lead to increased performance (Bandura, 1986, p. 391; Griffie, 1997a; Griffie & Templin, 1998). Inefficacious people avoid challenging goals that they fear will lead to negative outcomes. As a result, they do not perform as well.

Other Self-Phenomena

Self-efficacy is not exactly the same as other self-phenomena such as self-concept, self-esteem, confidence, and self-confidence (Ellis, 1990; Griffie, 1997b; Heyde, 1979; Larsen-Freeman & Long, 1991; Shavelson, Hubner & Stanton, 1976; Templin, 1995; Yule, Yanz & Tsuda, 1985), although some studies of self-efficacy mix it with these other self-phenomena (Huang & Chang, 1996; Mikulecky, Lloyd & Huang, 1996). Self-efficacy researchers specify five features that other self-phenomena researchers include only in part or not at all: (1) judgment of capabilities; (2) multiple dimensions; (3) contexts; (4) mastery-criterion; and (5) measurements taken before participants perform the task (Zimmerman, 1995). These are introduced below.

First, although self-efficacy is used as a judgment of capabilities (how well people believe they can do something), measures of other self-phenomena are often used as judgments of personal qualities (how well people feel about themselves). Second, self-efficacy researchers include multiple dimensions of research participants. Learners may believe they can introduce themselves orally, but they may not believe they can write a 50-word self-introduction. Other self-phenomena researchers do not always include multiple dimensions.

Third, self-efficacy researchers examine judgments of capabilities in various contexts. For example, learners may think they can introduce themselves in the context of a classroom of non-native English-speaking students, but they may think they cannot introduce themselves in a classroom of native English-speaking students. Although the task is the
same, the context is different. Other self-phenomena researchers do not depend on context.

Fourth, while self-efficacy is based on mastery criteria, other self-phenomena are usually based on normative criteria. Self-efficacy researchers specify how well learners believe they can accomplish tasks. Other self-phenomena researchers usually compare what learners feel about themselves in comparison with what other learners feel about themselves—a method that includes no direct measurement of what learners think they can actually do.

Finally, self-efficacy researchers need to measure self-efficacy before learners actually perform their tasks. Other self-phenomena researchers measure the self-phenomenon before the task, after the task, or without performance of the task at all. If researchers measure their self-phenomena after the task, or do not require participants to perform the task at all, they can predict nothing.

**Self-Efficacy Areas**

Other self-phenomena researchers have also been largely unsuccessful in predicting human behavior, whereas self-efficacy researchers have been widely successful. Researchers have successfully studied self-efficacy in a variety of areas that include, but are not limited to, academic achievement (Lee & Bobko, 1994; Lent, Brown & Larkin, 1984; Wood & Locke, 1987; Zimmerman, 1995), career choice and development (Hackett, 1995; Matsui, Ikeda & Ohnishi, 1989; Matsui & Tsukamoto, 1991), and health (Schwarzer & Fuchs, 1995).

Psychology and management researchers have repeatedly predicted that students with high self-efficacy attain higher grade point averages than students with low self-efficacy. Similarly, as students finish school, those with high self-efficacy in career pursuits and personal health experience more success in their career pursuits and health than those with low self-efficacy.

**Predicting L2 Learner Grades**

In studies attempting to predict L2 learners' grades in ESL settings, applied linguists recommend exploring factors such as motivation, personality, attitudes, previous knowledge, and previous academic performance to predict academic achievement (Graham, 1987; Light, Xu & Mossop, 1987; Patkowski, 1991). Even though psychology and management researchers have predicted academic success from self-efficacy measurements, applied linguists have not explored self-efficacy measurements as a way to predict academic achievement in language classes.
Statement of Purpose

The purpose of this exploratory research is to see if high self-efficacy students will achieve significantly higher grades than low self-efficacy students in an L2 learning class.

Method

Participants

The 74 participants in this study were tenth grade Japanese nationals in an urban high school ranked eighth out of nine high schools in its area in Kanagawa Prefecture. Students were enrolled in English I, which focuses predominantly on grammar-translation with some oral/aural instruction. There were 35 females and 39 males, ranging in age from 15-17. Students were in two intact classes instructed by the same teacher. All students participated by filling out a research questionnaire (see Appendix) after they had taken their first semester midterm exam, but before they received the results of the exam. This was done so participants would have feedback about the course, but would not base their responses only on grades (Wood & Locke, 1987). No language proficiency scores were available for these students.

Instrument

Considering the low level of the participants' high school and teachers' observations that previous students had poor English skills, the self-efficacy instrument was created in Japanese so students could fully understand the questionnaire. Japanese native speakers (fluent in English) and a non-native Japanese speaker (native English speaker) created the questionnaire in Japanese then translated it into English for non-Japanese readers (see Appendix). Contact the author for the Japanese original.

The self-efficacy measurement was adapted from Locke and Latham's (1990, p. 348) instrument, a composite of self-efficacy magnitude and strength. Magnitude has been used to measure the differing levels that subjects believe they can perform in a given domain. In the domain of academic achievement in an L2 class, this study asks students whether or not they believe they can achieve the following grades in their English class: F-, F, D-, D, C-, C, B-, B, A-, A. It may seem that measuring ten levels of academic achievement (F- to A) is overkill. However, measuring one level (whether or not students believe they can achieve As) gives no information about the differences between students who only believe they can achieve other levels (Bs, Cs, etc.). The self-efficacy
magnitude (see Appendix) shown in the left column, was obtained by asking students to answer yes or no if they could attain specific grades (F- to A). All data were entered into a ClarisWorks 4.0 (ClarisWorks Corp., 1994) spreadsheet and analyzed using Statview 4.5 (Abacus Concepts, 1995). The magnitude was then calculated by adding the total number of yes answers divided by the total number of items (10). Self-efficacy magnitude is the second most common self-efficacy measure in psychology and management research (Lee & Bobko, 1994). The most popular self-efficacy measure is self-efficacy strength (Bandura & Wood, 1989; Lee & Bobko, 1994; Matsui & Tsukamoto, 1991). People do not only differ in the levels of their efficacy beliefs (magnitude), but also differ in the strength of their efficacy beliefs:

Weak efficacy beliefs are easily negated by disconfirming experiences, whereas people who have a tenacious belief in their capabilities will persevere in their efforts despite innumerable difficulties and obstacles. They are not easily overwhelmed by adversity (Bandura, 1997, p. 43).

The questionnaire in the Appendix shows strength in the right column: Students rated their degree of confidence (0-100%) in attaining each grade level (F- to A). Strength was then calculated by adding the scores and dividing them by the total number of items (10).

Rather than using magnitude and strength scores independent of each other, Lee & Bobko (1994) recommend combining magnitude and strength scores for stronger predictive validity. The composite is calculated by adding the raw self-efficacy strength for grade levels that students answered yes to. Self-efficacy strength for grades answered no to are excluded. Fewer researchers (Gist, Schwoerer & Rosen, 1989; McAuley, Wraith & Duncan, 1991) use the composite self-efficacy instrument.

Table 1 shows the results of one student's questionnaire. This student wrote that, yes (magnitude), she thought she could score an F- in the English class for a final grade. This student was 100% confident (strength) about this. This student thought she could not score an F in the class. The student's confidence in scoring an F was 50%. The student thought she could not score anything higher and had no confidence in attaining any higher grade. The researcher divided the number of yes scores (1) by the number of levels (10) for the student's magnitude score (.10). Then the researcher added all of the strength scores (.15 + .00 + .00, etc.) and divided by 10 for the student's strength score (.15). Finally, the researcher added all of the strength scores for yes answers (1.00 for F-). All strength scores for no answers (.50 for F, etc.) were excluded. This student's scores are the lowest scores in Table 2 for magnitude, strength, and composite. Although not observable from the data presented here, this student's final English grade was F (F=2).
Table 1: One Student’s Magnitude, Strength, & Composite Scores

<table>
<thead>
<tr>
<th>Grade (Yes/No)</th>
<th>Magnitude</th>
<th>Strength (.0-1.00 Confidence)</th>
<th>Composite (Strength of Yes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F- Yes</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>F No</td>
<td>.50</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>D- No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>D No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>C- No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>C No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>B- No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>B No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>A- No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>A No</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
</tbody>
</table>

Scores .10 (average) .15 (average) 1.00 (sum)

Grades were determined by the teacher of the two classes by averaging grades for three semesters. These included grades for exams, assignments (in and out of class), and attendance and were represented on a scale of 1-10, the lowest score being 1 (F-) and the highest score being 10 (A).

Reliability of the Instrument

The reliability of the self-efficacy scores and grades were calculated using Cronbach’s alpha and are reported in Table 2 below. The two subsections, magnitude and strength, and the composite of the questionnaire are .91, .98, and .96, respectively. The reliability of grades could not be determined because the necessary data were not available to the researcher.

During class the teacher passed out the questionnaire and gave students 10-15 minutes to fill it out. She suggested the students would probably answer yes with 100% confidence for the first question, since it is impossible to score lower than an F-. She did not recommend answers for any of the other questions.

After the students finished the questionnaires, the teacher collected them and sealed them in an envelope that she handed to the researcher after class. The teacher never saw the results of the questionnaires. At the end of the school year, the teacher gave her students’ grades to the researcher.
Table 2: Descriptive Statistics for Self-Efficacy Scores and Grades

<table>
<thead>
<tr>
<th></th>
<th>Subtests</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Magnitude</td>
<td>Strength</td>
<td>Composite</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>74.00</td>
<td>74.00</td>
<td>74.00</td>
<td>74.00</td>
</tr>
<tr>
<td><strong>k</strong></td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>3.00</td>
</tr>
<tr>
<td><strong>M</strong></td>
<td>.53</td>
<td>.50</td>
<td>4.48</td>
<td>6.47</td>
</tr>
<tr>
<td><strong>Mode</strong></td>
<td>.50</td>
<td>.66</td>
<td>5.00</td>
<td>6.00</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>.50</td>
<td>.49</td>
<td>4.30</td>
<td>6.00</td>
</tr>
<tr>
<td><strong>Midpoint</strong></td>
<td>.55</td>
<td>.55</td>
<td>5.30</td>
<td>5.50</td>
</tr>
<tr>
<td><strong>Low-High</strong></td>
<td>.10 - 1.0</td>
<td>.5 - .96</td>
<td>1.0 - 9.6</td>
<td>1.0 - 10</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>1.90</td>
<td>1.81</td>
<td>9.60</td>
<td>10.00</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>.17</td>
<td>.16</td>
<td>1.60</td>
<td>1.98</td>
</tr>
<tr>
<td><strong>Chronbach’s Alpha</strong></td>
<td>.91</td>
<td>.98</td>
<td>.96</td>
<td>*</td>
</tr>
</tbody>
</table>

*unavailable

Statistical Analysis

To analyze the data, descriptive statistics were calculated for the self-efficacy scores and grades (Table 2). The self-efficacy scores and grades have similar means, modes, medians, and midpoints. Differences were measured by a paired t-test, with an alpha level of .05.

Table 3: Low and High Self-Efficacy Students’ Grades

<table>
<thead>
<tr>
<th></th>
<th>Groups</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>37.00</td>
<td>37.00</td>
<td></td>
</tr>
<tr>
<td><strong>k</strong></td>
<td>3.00</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td><strong>M</strong></td>
<td>5.89</td>
<td>7.05</td>
<td></td>
</tr>
<tr>
<td><strong>Mode</strong></td>
<td>6.00</td>
<td>7.00</td>
<td></td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>6.00</td>
<td>7.00</td>
<td></td>
</tr>
<tr>
<td><strong>Midpoint</strong></td>
<td>5.50</td>
<td>6.50</td>
<td></td>
</tr>
<tr>
<td><strong>Low-High</strong></td>
<td>1 - 10</td>
<td>3 - 10</td>
<td></td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>10.00</td>
<td>8.00</td>
<td></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>1.89</td>
<td>1.92</td>
<td></td>
</tr>
<tr>
<td><strong>SD squared</strong></td>
<td>3.59</td>
<td>3.71</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Results of T-test Comparing Grades of Low & High Self-Efficacy Students

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean Difference</th>
<th>df</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low, High</td>
<td>-1.16</td>
<td>36</td>
<td>-2.85*</td>
</tr>
</tbody>
</table>

*p < .05

Results

In order to compare the grades of low self-efficacy students with the grades of high self-efficacy students, the independent variable of this study was defined as the student's grade and the total number of participants, 74, was divided into halves. Those students who scored in the lower half on the self-efficacy composite were designated as the low self-efficacy group and students scoring in the upper half were designated as the high self-efficacy group. The descriptive statistics are given in Table 3.

Since both the low and high self-efficacy groups meet the assumptions of grouping, continuous data, normal distributions, and equal variance for a t-test, a one-tailed t-test was selected to compare group means (see Table 4).

As shown, the difference between the grades of low self-efficacy and high self-efficacy students was significant at p < .05.

Discussion

This pilot study suggests that high self-efficacy students achieve significantly higher grades than low self-efficacy students in an L2 classroom. From the beginning of the school year, low self-efficacy learners believe they cannot succeed academically and thus remain cut off from higher achievement throughout the year. This result is in agreement with self-efficacy research in psychology and management that shows low self-efficacy learners decrease attention, effort, persistence, and strategies for achieving, and they avoid challenging goals. While this researcher has observed that some students only exhibit low self-efficacy in language learning classes (e.g., they exhibit high self-efficacy in math, extracurricular activities, etc.), other students exhibit low appraisals of their capabilities across many of their school activities—a sign that these students may be in particular need of help.
Someone might argue that self-efficacy is just sound self-knowledge—people already know what they can and cannot do. But people do not always know what they can and cannot do (for more on the discordance between efficacy judgment and action, see Bandura, 1997, pp. 61-78). In dangerous situations where mistakes can be fatal, people kill themselves by overestimating their capabilities. However, in less dangerous situations, underestimating one's capabilities can lead to regret; "Educational opportunities forsaken, valued careers not pursued, interpersonal relationships not cultivated, risks not taken, and failures to exercise a stronger hand in shaping one's life course" (Bandura, 1997, p. 71).

Bandura (1995) cites research that shows four ways people can raise their self-efficacy. The first way is through *enactive mastery experience*. Learners need opportunities to experience success in L2 learning classrooms. Also, instead of measuring students' mastery using norm-referenced tests (NRTs) that only allow about 2% of the students to receive As, teachers should use criterion-referenced tests (CRTs) in their classrooms. Criterion-referenced tests allow 100% of the students to receive As and measure mastery of the coursework (Brown, 1996).

Second, learners can increase their self-efficacy through *vicarious experience*. When learners see their peers—whom they judge to be of similar L2 proficiency—fail, learners expect to fail. In contrast, learners who see their equals succeed believe they can succeed, too. Also, when Japanese teachers of English speak English, students believe that they can speak English, too.

*Verbal persuasion* is a third way learners can increase their self-efficacy. People can be persuaded verbally that they can succeed. Bandura (1995) explains,

> Successful efficacy builders do more than convey positive appraisals. In addition to raising people's beliefs in their capabilities, they structure situations for them in ways that bring success and avoid placing people in situations prematurely where they are likely to fail often. They encourage individuals to measure their success in terms of self-improvement rather than by triumphs over others. (p. 4)

Depending on what messages teachers send to their students, teachers can influence whether students have high or low self-efficacy.

Fourth, *physiological and affective states* affect learners' beliefs in their capabilities. Learners need to understand how to interpret feelings of arousal as positive, and learners need to be healthy. For example, before speaking in an L2, if students interpret their increased heartbeats, faster breathing, and higher perspiration as debilitating, they will lower their self-efficacy. Students with a positive interpretation will use the arousal to energize their
performance. In addition, students need to get proper amounts of rest, eat a balanced diet, exercise regularly, etc. (For creating a self-efficacy syllabus in an EFL classroom, see Templin, in press.)

Although this study indicates that learners with high self-efficacy perform higher academically, it does not necessarily show that learners will successfully acquire the L2 studied. One difficulty with measuring L2 acquisition in Japanese academic institutions is that reliable and valid L2 proficiency measurements are rare. This researcher has advised and participated in language testing at the high school and university level, including administration of the Ministry of Education-endorsed eiken (tests produced by STEP, the Society for Testing English Proficiency). Reliable and valid testing is the exception rather than the norm (see articles in Brown & Yamashita, 1995), yet such measurements are needed so researchers can find out how much of the L2 learners actually acquire.

Also, using a composite of self-efficacy magnitude and strength scores is cumbersome to calculate. In this study, calculating strength alone seemed just as satisfactory as calculating a composite measure. Bandura (1997), says that calculating strength alone "provides essentially the same information and is easier and more convenient to calculate" (p. 44).

In future studies of academic achievement in L2 classrooms, it is suggested that researchers investigate self-efficacy instruments that measure the other dimensions of academic achievement such as concentration, memorization, and note-taking (Lee & Babko, 1994; Wood & Locke, 1987).

Acknowledgments
The author thanks J.D. Brown, Dale T. Griffie, Nicholas O. Jungheim, Cynthia Lee, and Tamao Matsui for their comments regarding this manuscript. Correspondence should be addressed to Stephen A. Templin, Meio University, International Cultural Studies Department, 1220-1 Bitumata, Nago-shi, Okinawa, Japan 905-0005. E-mail: steve@ics.meio-u.ac.jp. Work fax: 0980-52-4640.

Stephen A. Templin is the author of Communicative Tool Box for Japanese Students (Seido Language Institute). His articles have appeared in JALT Journal, The Language Teacher, TESL Reporter, and other publications.

References
Cambridge: Cambridge University Press.


PERSPECTIVES

chology, 31, 356-362.

(Received February 6, 1998; revised October 15, 1998)
Appendix: Self-Efficacy Questionnaire (English Version)

__Year__ __Class__ __ID__ Male_ Female_ Name____________________

(Your teacher will not look at this, and your answers will not affect your grades.)

In this class (for your final grade),

Do you think you can score an F-?
Yes_ No_

Do you think you can score an F?
Yes_ No_

Do you think you can score a D-?
Yes_ No_

Do you think you can score a D?
Yes_ No_

Do you think you can score a C-?
Yes_ No_

Do you think you can score a C?
Yes_ No_

Do you think you can score a B-?
Yes_ No_

Do you think you can score a B?
Yes_ No_

Do you think you can score an A-?
Yes_ No_

Do you think you can score an A?
Yes_ No_

How much confidence do you have that——

You can score an F-? (0% - 100%)________
You can score an F? (0% - 100%)________
You can score a D-? (0% - 100%)________
You can score a D? (0% - 100%)________
You can score a C-? (0% - 100%)________
You can score a C? (0% - 100%)________

You can score a B-? (0% - 100%)________
You can score a B? (0% - 100%)________
You can score an A-? (0% - 100%)________
You can score an A? (0% - 100%)________

Note: The original Japanese questionnaire can be obtained by contacting the author.