Predicting English Language Attainments in Higher Education: Self-efficacy Scales vs. Past Performances

Matthew Rudd
Doctoral graduate (Ed.D), Philippine Christian University
matthew.rudd@hotmail.com

Abstract—This paper aims to empirically test Bandura’s (1977; 1986) claims that the beliefs people hold in their attainment potential predict standards of achievement more accurately than records of past performance, as the latter represent quiescent events unable to account for change. To investigate Bandura’s theoretical ideology, this quantitative research study thoroughly tested the accuracy of self-efficacy scales by comparing the grades 111 third-year university business students believed themselves capable of achieving vis-à-vis subsequent end-of-year English language scores. In relation to past performances, the present study compared students’ previous second-year scores with those attained at the end of their current third year. This helped to determine whether students’ future performances can be computed more accurately as a function of their personal efficacy beliefs rather than by assumptions based on past performances. The results of this experiment were mixed. Initial t-test calculations inferred that there was no significant difference between prior stated beliefs of self-efficacy and end-of-third-year English scores (English Y3 = 63%; self-efficacy = 64.23%, p <0.05, not sig). In the same vein, students’ prior second-year scores were significantly different to those achieved in their third year (English Y3 = 63%; past performance Y2 = 56.64%, p <0.05, sig); suggesting that self-efficacy beliefs more reliably project future performances than past exploits. Nevertheless, while self-efficacy scales proved to be accurate in general terms, when measuring on a more task-specific level (such as speaking) or sub-group level (namely student ability), it transpired that such data deriving from self-efficacy scales became increasingly inconsistent. These confounding implications are discussed at length in this paper.

Key Words: self-efficacy; past performance; accuracy; achievement; belief.
Despite extensive research findings that claim strong correlations between past performances and future attainments (discussed later in this paper), Bandura (1977; 1986) nevertheless posits that self-efficacy measures constitute the most reliable method to inform of people’s academic potential. To investigate this claim, the current quantitative research study looks to answer two fundamental research questions: (1) to investigate the correlation between students’ (N = 111) end-of-third-year English scores in contrast to their prior stated beliefs of self-efficacy; and, (2) to compare students’ past performances in English (second year) with those attained successively in their third year. Both lines of inquiry will clarify and compare the predictive accuracy of both variables; self-efficacy and past performance. Further cross-sectional analyses will also look to examine the possible interaction of gender, student ability levels and task-specificity in the aforementioned parameter.

Before such discoveries can be arrived at, it is necessary to summarise the key tenets of self-efficacy as a belief system that explain and predict psychological changes. Self-efficacy beliefs are summarised as “the level of conviction in one’s capabilities to successfully execute the behaviour required to produce outcomes” (Bandura, 1977). According to Bandura (1977; 1989; 2009) it is from the constellations of information contained in environmental events that contribute to forming people’s level of self-efficacy, which originate from four main sources: (1) enactive mastery (past performance accomplishments), (2) vicarious experiences (observing the performances of others), (3) exhortative influences (verbal persuasion from others), and (4) emotive sources (physiological states).

The most significant contributor to self-efficacy development is through mastery experiences (Bandura, 1977; 1994), as they provide direct evidence of achievement; and it is through experiencing gradual successes that mastery expectations are raised (Bandura, 1977). Vicarious experiences, encapsulated as the observed successes and failures of others, do not provide an authentic experiential base (Bandura, 1977), rendering this nature of information less dependable than personal mastery (Bandura, 1994). Although we do learn from others, if the model from which the information derives is not perceived to be similar to the observer, or, the observer completes the demonstration with noticeable ease, the effect of modelling on self-efficacy beliefs tends to be limited (Bandura et al., 1996).

Thirdly, persuasive influences from peers can drastically reinforce or undermine self-efficacy beliefs; regrettably negative talk is more influential than positive persuasion (Bandura, 1994). This is believed to be due to the fact that our minds are highly attuned to threat responses; actions that compromise our status and affects our level of functioning (Rock, 2008). Moreover, if the nature of persuasion is perceived to be unrealistic, such ephemeral boosts in efficacy are soon disconfirmed by subsequent sub-standard performances (Bandura, 1994).

The fourth source of efficacy information stems from frame of mind and emotional states. Those plagued by doubts and depressive occurrences suffer from undermined personal beliefs, whereas,
positive thinking enhances perceptions of personal efficacy, equipping individuals with stronger beliefs to face and overcome difficult challenges; “people who have a high sense of efficacy are likely to view their state of affective arousal as an energizing facilitator of performance” (Bandura, 1994).

Efficacy information conveyed by environmental events is interpreted and processed by the individual by four mediating self-efficacy activated processes which affect and regulate behaviour (Bandura, 1977; 1994). Firstly, cognitive processes shape courses of action people (pre)conceive, and the envisaged outcomes that may result (Bandura, 1994). Those with a high sense of self-efficacy visualise successful scenarios, strategize purposeful courses of action that successfully accomplish challenging tasks. Fruitful completion of these ventures conveys confirmatory evidence of enhanced competence, subsequently reinforcing efficacy beliefs (Bandura, 1977).

Secondly, beliefs of personal efficacy regulate motivation through their impact on aspirations as well as the determination invested to see to those goals and (Bandura, 2009). More efficacious individuals set ambitious long-term goals and organise courses of action to bring these valued futures into fruition, especially through accumulative sub-goal attainments (Bandura, 1994; Bandura & Schunk, 1981). Thirdly, self-efficacy plays a central role in managing anxiety arousal and sustaining staying power in the face of adversity (Bandura, 1994). It is the development of coping skills that enhances efficacious thinking, which minimises the interference of affective processes on one’s level of functioning (Bandura, 1994).

Lastly, beliefs of personal efficacy also affect one’s selection processes, and can shape life trajectories influencing the type of academic and vocational careers people choose (Bandura, 1994). Pajares (2003) concluded in a detailed review of literature that judgments of personal efficacy exert decisive influence on students’ selection behaviour and the nature of the pursuits they choose, and it is efficacy beliefs that govern the level and intensity of effort invested. More profoundly, Bandura (2009b) categorically emphasises the centrality of a belief system as a quintessential concept to govern motivation, and that “rational models of motivation and decision making that exclude perceived self-efficacy sacrifice explanatory and predictive power”.

The fundamental role of self-efficacy influences all academic domains (Pajares & Schunk, 2001; Zimmerman, Bandura, & Martinez-Pons, 1992), and, the development of self-efficacy beliefs and competencies may concurrently develop across unrelated spheres of activity; as many domains are governed by similar underlying sub-skills (Bandura, 1997; 2006). Likewise, perceived cognitive inefficacy has been linked to lowered intellectual performances (Lachman, Steinberg, & Trotter, 1987). Pintrich and De Groot (1990) clarify that higher self-efficacy beliefs lead to increased use of metacognitive strategies, which successively improve the standard of achievements, but also caution that students must have both ‘the will and the skill’ to be successful in classrooms (p.38).

**Literature Review**

For a thorough review of literature, this section will analyse findings from
research studies which have investigated the utility of both past performances and self-efficacy scales in predicting academic attainment potential.

**The reliability of past performances**

A large-scale study in the Netherlands (Med Educ. 2015) analysed the correlation between past performances (in the form of university entrance exams) and end-of-first-year grades among 2,357 under-graduate medical students from six cohorts between 2002-2007. The results confirmed a strong link between the two sets of results, as pre-university academic standards were closely associated with academic achievements at the end of the first year, and correspondingly, end-of-first-year GPAs were closely correlative to performance standards at the end of the 4th year.

Similarly in the Netherlands, a paper published in PLOS ONE (2016) studied whether students’ first-year academic performance could be accurately mapped out using university entrance results, and or by former high school attainments. During the academic year 2013-2014 data derived from 652 applicants, (mean age: 20 years), revealed that university admission testing was the most accurate predictor of end-of-first-year grades ($r = .56$). Prior high school attainments were a less reliable indicator, as many students were from overseas, originating from heterogeneous education systems. However, the Dutch students’ high-school attainments were reasonably parallel to end-of-first-year university achievements ($r = .52$).

In concert with the findings reported from the two Dutch research papers, the European Journal of Dental Education (2018) also claimed testimony to the correlative strength of past performances. The study in question comprised 116 oral health students at the Auckland University of Technology, comparing current attainments in bioscience with previous GPA scores. The study emphatically concluded that prior academic performance was a statistically significant predictor of academic behaviour ($r = 0.641$).

**The reliability of self-efficacy scales**

A detailed report was published in the Journal of Dental Education (March, 2019) which studied the relationship between students’ self-efficacy and academic performance at a dental school at the Universidad Cooperativa de Colombia. This cross-sectional study was originally conducted in 2016 with dental students from all five year groups, whose measurements of self-efficacy were taken using the psychometric properties of the Academic Behaviors Self-Efficacy Scale (ABSS).

Current grade point average was used to measure the academic performance of 320 students, and in contrast to which, the results inferred a timid association between the former and self-efficacy data ($\rho = 0.259; p<0.001$). That said, a more positive correlation was observed among female students ($\rho = 0.361; p<0.001$), and intriguingly, those in low socioeconomic status strata (1-2: $\rho = 0.310; p<0.001$), as well as those with heavier workloads ($\rho = 0.306; p = 0.001$). The interesting observation from this study was that females declared more congruent levels of self-efficacy than male students, which is in agreement with Britner and Pajares’ (2001) findings in a separate educational context, concluding from a
sample of 262 seventh grade African American pupils that girls displayed higher levels of self-efficacy, and subsequently scored higher academically on science tests.

Pajares, Miller and Johnson (1999) had also found from their study of 363 elementary students (grades three, four and five) that female elementary students were also more self-efficacious and higher achievers than their male counterparts in writing efficacy; and specifically declared themselves to be more apt than boys. This was also verified in Pajares and Valiante’s (1999) middle-school based study encompassing a larger sample population (N = 742), which extended the scope to cover students in grades six, seven and eight. Based on his findings, Pajares (2003) later summarised that female students tend to express greater confidence in language arts, and also reached higher levels of attainment.

Self-efficacy beliefs have also been found to extend to kinaesthetic-based spheres. McCormick and McPherson’s (2003) research showed that out of the 332 young instrumentalists who were preparing for Trinity College London performance Examinations for music, self-efficacy was found to be the best predictor of performance (more so than study behaviour). The authors also confirmed this to be the case in a follow up study carried out in 2006, consisting of 686 students (grades one to eight) sitting Australian Music Examinations Board (AMEB) performance examinations. The results highlighted that 100-point efficacy assessments measuring 5 specific test areas proved to be a highly accurate predictor.

The implementation of self-efficacy scales

It is worth pointing out that the four aforementioned studies (except the Colombia based paper) measured efficacy levels on a sub-domain or task specific measure, and also used 100-point response format, helping to yielded informative results. The Colombia based study on the other hand attempted to deploy self-efficacy measures to predict general levels of performance in terms of GPA, which revealed differing empirical qualities. This measure was too broad and not subject specific, as this measure includes a plethora of separate subjects and sub domains.

When constructing self-efficacy scales, items being measured must also closely correspond with the domain being investigated (Pajares, 1996; Bandura, 1995). Bandura (2006) emphasises that there is no all-purpose measure of efficacy beliefs, as generic measures may have little or no relevance to the domain of functioning (p.307). Domain-specific measures are more explanatory than general measures, as the latter "violate the basic assumption of the multidimensionality of self-efficacy beliefs … nor do they have much predictive utility" (Bandura, 1997, p.48). Pajares (1996) adds further scrutiny, proposing that task-specific measures are more accurate on the basis that mastery of distinct skills is essential to perform highly in any given domain; for instance, mathematics requires efficacy in calculations, interpreting fractions, and working out sums of division or multiplications, and so forth.
A further implication is that efficacy measures are best squared against test areas (see McCormick & McPherson, 2003; 2006), whereas in the Colombian dental study questions were not directly relevant to academic performance, but more so focused more on classroom behaviour and communication skills, such as: confidence in expressing ideas, engaging in dialogue with the professor, working with peers, punctuality, attendance levels, submitting homework on time, and, preparing for exams. It could be argued that these elements are the constituent sources of self-efficacy beliefs; nevertheless, these rubrics cannot clearly be measured to predict academic achievements.

Having analysed the accuracy of past performances, the theoretical framework of self-efficacy and the respective practical implications, the three hypotheses formulated are as follows:

(1) Measures of self-efficacy beliefs are expected to correlate closely with final grade outcomes (Bandura, 1977).
(2) Efficacy Beliefs are assumed to correlate more closely with end-of-year grades than students’ prior performances (Bandura, 1986).
(3) a.) Girls’ declared levels of efficacy are hypothesised to correlate more accurately than those disclosed by their male counterparts (Britner & Pajares, 2001; Pajares, 2006).
b.) More specific measures of self-efficacy are presumed to yield greater accuracy (Pajares, 1996).

Methodology
Research design, participants and locale of study

Following executive approval, this study took place during academic year of 2018-19 at a university of technology situated in the outskirts of Bangkok. The administration team kindly provided all data for past grades as well as the confirmation of the official end-of-year grades for the current year (2018-2019). The central focus of the study was to compare the correlative validity of perceived efficacy beliefs vis-à-vis end-of-third-year grades in English language; and, whether self-efficacy measures correlate more closely with end-of-third-year grades than students’ performances from the year before.

The prescribed method for constructing efficacy scales is to provide a 100-point continuum that questions what respondents can do rather than what they will do, seeing that can denotes a judgment of capability, whereas will refers to a statement of intent; while perceived self-efficacy is a central determinant of intent, the two constructs are separate (Bandura, 2006).

Self-efficacy scales present items with different levels of task demands, and respondents rate the strength of their belief in their ability to successfully perform the corresponding task on a 100-point scale (Bandura, 2006); ranging from uncertainty, intermediate values of certainty, to great certainty (blocked in units of 10) for each individual question (Bandura, 1977; 2006). This constitutes a more accurate means than restricted continuums because respondents tend to avoid extreme positions opting for the middle value, thus
more mid-way options allow for a greater variety in responses (Bandura, 2006, p.312). In accordance with this, Pajares, Hartley, and Valiante (2001) found from further research in a middle school context (N = 497) that the 100-point scale offered more psychometrical strength than the 6-point Likert scale. As seen below, the 100-point efficacy scale used in the experiment naturally reflects the grading criteria, which is also scored out of one hundred (in terms of a percentage).

**Table 2. Conversion of students’ percentage scores into letter grades.**

<table>
<thead>
<tr>
<th>Percentage Score</th>
<th>Letter Equivalent</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>90</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>B+</td>
<td>B</td>
</tr>
<tr>
<td>70</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>C+</td>
<td>C</td>
</tr>
<tr>
<td>50</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>D+</td>
<td>D</td>
</tr>
<tr>
<td>30</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

Students were asked to score their degree of confidence in four test areas (see Appendix). Firstly, on a domain-specific level, respondents were asked to highlight the overall grade in English they believed themselves capable of achieving. In terms of the two sub-domain questions (reading related), students marked the grade outcomes they felt they could attain in both the mid-term and end-term tests. And, in relation to fourth task-specific measure, respondents scored their degree of certainty in their attainment potential in a prescribed speaking test, which simulated a job interview situation.

Acknowledging the fact that students need to be familiar with task demands (Bandura, 1977; 2006), 111 third-year business students (aged 20-21) were selected for this study given their level of familiarity with the university’s syllabus, course materials and teaching methodologies. The demographics were also well-balanced (females =59; males = 52), which opened the scope to analyse the interaction of the gender variable.

**Data collection and analysis**

English language courses comprise three main forms of assessments: mid-term exams (weighted at 20%), speaking tests (30%), and, end-of-term examinations (50%). The mid-term and end-term assessments consist primarily of reading comprehension related questions. Efficacy scales were implemented to ascertain students’ confidence levels in overall grade attainments (domain-related), mid-term and end-term assessments (reading: sub-domain), and a speaking test simulating a job interview (task-specific). It would be of additional interest to see whether the difference in the temporal distance between the mid-term tests (taken in week 8) and the end-term exams (week 16) would influence the relationship between efficacy declarations and attainment levels.

After data collation, the correlation between students’ end-of-third-year grade outcomes and their previously stated efficacy beliefs were compared using t-tests and Pearson’s Coefficient. The same instruments were utilised to examine the relationship between current end-of-third-year and prior end-of-second-year grade outcomes. Correlations from both tests were contrasted for analyses, and further cross-sectional analyses were also
conducted to ascertain the intervening role of gender, student ability grouping and task-specificity to help identify and elucidate any potential confounding factors.

Ethical considerations

According to Bandura (2006) self-efficacy scales must be conducted anonymously for reasons pertaining to privacy and data confidentiality. This also helps to improve the accuracy of data as students feel freer to respond honestly. Furthermore, only basic demographic information relating to gender was required in order to analyse the interaction of the gender variable within the context of the variables being studied. Although students also had the right to refrain from completing the questionnaire, all participants were willing to cooperate.

Findings

The central research question of this paper was to investigate the correlative strength between students’ end-of-year English language attainments and their prior stated levels of self-efficacy. The secondary question set out in this paper was whether students’ end-of-year performances correlated more closely with prior statements of self-efficacy or their past performances from the year before. Thirdly, further cross-sectional analyses of additional variables, such as gender, ability level and task-specificity may provide further explanatory-specific cues than a single general calculation.

Table 2: The correlation between self-efficacy beliefs and past performances (N = 111).

<table>
<thead>
<tr>
<th>English Score (Y3)</th>
<th>Variable</th>
<th>M %</th>
<th>P</th>
<th>p &lt; .005</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-efficacy</td>
<td>64.23</td>
<td>0.276</td>
<td>not sig</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Past Performance (Y2)</td>
<td>56.64</td>
<td>0.002</td>
<td>Sig</td>
<td>0.81</td>
</tr>
</tbody>
</table>

The t-tests in Table 2 indicate that self-efficacy measures more accurately mirrored end-of-third-year grade outcomes than did past second-year performances, supporting Bandura’s (1977; 1986) theoretical claims. This is reflected in the p value between efficacy beliefs and end-of-third-year attainments (p = 0.276; not sig), denoting that there is no significant difference between the two variables. Furthermore, the level of incongruence noted between students’ current end-of-third-year grades and those achieved in the second year were of significantly differing values (p = 0.002; sig). Despite this observation, past performances correlated very strongly with the students’ current level of achievement (r = 0.81). This correlation appeared to be stronger than those recorded in the university level studies discussed in prior literature, (which was noted between .56 and .64); whereas self-efficacy beliefs displayed a notably weaker relationship (r = 0.49). The mixed nature of the results complicates the legitimacy of the first hypotheses stated previously, that self-efficacy measures closely correlate with
future attainments (Bandura, 1977). The second hypothesis also falls under scrutiny, seeing that efficacy beliefs were expected to correlate with future attainments more so than past performances (Bandura, 1986). Looking at sub-groups, such as gender and task specificity, may help to identify the root cause of the inconsistencies noted in the r value.

**Table 3: The role of the gender variable in self-efficacy beliefs and performance outcomes.**

<table>
<thead>
<tr>
<th>Gender</th>
<th>n</th>
<th>English (Y3): M%</th>
<th>Variable</th>
<th>M%</th>
<th>P</th>
<th>p &lt; .005</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Self-efficacy</td>
<td>65.08</td>
<td>0.333</td>
<td>not sig</td>
<td>0.526</td>
</tr>
<tr>
<td>Females</td>
<td>59</td>
<td>63.8</td>
<td>Past Perf. (Y2)</td>
<td>58.41</td>
<td>0.045</td>
<td>sig</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Self-efficacy</td>
<td>63.27</td>
<td>0.341</td>
<td>not sig</td>
<td>0.437</td>
</tr>
<tr>
<td>Males</td>
<td>52</td>
<td>62.1</td>
<td>Past Perf. (Y2)</td>
<td>54.63</td>
<td>0.010</td>
<td>sig</td>
<td>0.794</td>
</tr>
</tbody>
</table>

As observed in Table 3, t-test results indicate that both female and male students had stated equally accurate levels of self-efficacy expectations and consequently achieved identical levels of performance; as noted in the identical p values (females, p = .333; males, p = .341). These findings are in disagreement with Britner and Pajares’ (2001) and Pajares, Miller and Johnson’s (1999) claims that girls demonstrate higher levels of self-efficacy, and thus achieve higher levels of academic performance (especially in language arts); consequently invalidating the third hypothesis stated in this paper. Moreover, across both gender categories, t-tests also insinuated that end-of-third-year English attainments were significantly higher than those recorded the previous year (females: p = 0.045, sig; males: p = 0.010, sig). This clearly manifests that both male and female students are capable of improving and that the past is not necessarily a guide to the future.

That said, the modest correlations between self-efficacy beliefs and end-of-third-year performances consolidates the allusion that there is a degree of inconsistency within the data; despite males and females being of a similar standard academically and stating similar strengths of self-efficacy beliefs.

**Table 4: Accuracy of self-efficacy and past performance according to ability tier.**
The data patterns that emerge in Table 4 begin to question the accuracy of self-efficacy measures in relation to student ability levels. The t-tests confirm that lower-performing students significantly overstate their attainment potential ($p < .001$, sig), while higher achieving students significantly understated their achievement capacity ($p < .001$, sig). The mid-tier group on the other hand stated efficacy beliefs parallel to their end-of-third-year performances ($p = .109$; not sig). Nonetheless, even among the mid-tier group, the results highlighted in Table 4 still point to a very weak correlation between declared efficacy beliefs and end-of-third-year performances ($r = 0.21$); still inferring high levels of variance within the data.

However, correlations between end-of-third-year performances and prior performances seemed indicate a stronger and more constant relationship, especially among the mid and upper-level sub groups ($r = .348$ and $.525$ respectively); despite significantly differing means. Remarkably, the reverse was noted among the lower-level sub group, whose data sets suggest a weaker correlation in spite of very similar means ($p = 3.49$; not sig). On a separate note, improvements noted among mid-tier and upper-tier students highlight that past performance are not necessarily a reliable guide to the future, and ability levels can evolve and develop. This also begins to imply that beliefs may be a driver for such improvements.

<table>
<thead>
<tr>
<th>Tier</th>
<th>n</th>
<th>M %</th>
<th>Variable</th>
<th>M %</th>
<th>P</th>
<th>p &lt;.005</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>37</td>
<td>83.56</td>
<td>Self-efficacy</td>
<td>71.35</td>
<td>&lt;.001</td>
<td>sig</td>
<td>0.348</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Past perf. (Y2)</td>
<td>72.94</td>
<td>&lt;.001</td>
<td>sig</td>
<td>0.693</td>
</tr>
<tr>
<td>Mid</td>
<td>37</td>
<td>61.7</td>
<td>Self-efficacy</td>
<td>64.05</td>
<td>0.109</td>
<td>not sig</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Past perf. (Y2)</td>
<td>52.6</td>
<td>&lt;.001</td>
<td>sig</td>
<td>0.525</td>
</tr>
<tr>
<td>Lower</td>
<td>37</td>
<td>43.7</td>
<td>Self-efficacy</td>
<td>57.4</td>
<td>&lt;.001</td>
<td>sig</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Past perf. (Y2)</td>
<td>44.37</td>
<td>0.349</td>
<td>not sig</td>
<td>0.157</td>
</tr>
</tbody>
</table>

**Table 5:** Analysing the accuracy of self-efficacy scales according to domain-specificity.
The results in Table 5 display further inconsistencies regarding the accuracy of specific efficacy measures. On a sub-domain level (reading comprehension), the results showed high levels of congruence between efficacy beliefs and mid-term as well as final exam results (mid-term: $p = 0.133$, not sig); (final: $p = 0.256$, not sig). Nevertheless, proposals that task-specific measures offer greater reliability (see Pajares, 1996) also seem to have been put into question. The mean values of stated self-efficacy beliefs and end-of-year speaking scores (simulated job interview) were grossly incommensurate (<.001, sig) despite inferring a relatively firm correlation ($r = .536$). Based on this observation, these results do not corroborate the fourth hypothesis put forward in this paper. Regrettably, task-specific and sub-domain-specific data pertaining to past performances were not available given that students prior scores are archived as a single grade, and itemised data were not retrievable. This constitutes one of the limitations of the current study.

### Discussion

The findings in this study have brought to light a set of mixed results surrounding the predictive correlation of personal efficacy beliefs and past performance. As a result, a number of confounding factors warrant further discussion, which include instrumentation, gender, student ability levels and task-specificity.

#### Instrumentation

As hitherto discussed, t-tests inferred that end-of-third-year grade outcomes were relatively commensurate with prior stated beliefs of self-efficacy, while Pearson's Coefficient calculations implied that the two sets of results were at variance; and that end-of-third-year grades were more closely commensurate with students’ past attainments from the previous academic year.

One plausible explanation to account for the later trend was that a universal level of incremental improvement was achieved by most students, which would logically reflect in a stronger $r$ value. The most useful aspect of the Pearson’s Coefficient instrument was that it helped to signal discrepancies between final third-year grade outcomes and previously stated self-efficacy beliefs, which lead to the discovery that self-efficacy measures were inconsistent in more specific measures; prompting further

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Variable</th>
<th>M %</th>
<th>P</th>
<th>$p &lt; .005$</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>English (Y3)</td>
<td>Mid-term</td>
<td>56.85</td>
<td>0.133</td>
<td>not sig</td>
<td>0.336</td>
</tr>
<tr>
<td>(sub domain)</td>
<td>Self-efficacy</td>
<td>59.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>English (Y3)</td>
<td>58.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final exam</td>
<td>Mid-term</td>
<td>0.256</td>
<td></td>
<td>not sig</td>
<td>0.538</td>
</tr>
<tr>
<td>(sub domain)</td>
<td>Self-efficacy</td>
<td>59.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>English (Y3)</td>
<td>68.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>Mid-term</td>
<td>&lt;.001</td>
<td>sig</td>
<td></td>
<td>0.536</td>
</tr>
<tr>
<td>(task specific)</td>
<td>Self-efficacy</td>
<td>56.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
investigation. The utilisation of both instruments played an essential role to highlight these discoveries, offering more statistical depth. The absence of both statistical calculations in prior literature did not help to clarify whether the standards of achievement were significantly different or similar, merely that there was a correlation.

**Gender**

The results suggested that male and female students’ efficacy beliefs relative to end-of-third-year language performances displayed similar trends according to both t-tests and coefficient analysis. Such empirical evidence did not readily validate prior claims, and disproved the third hypothesis of this paper, that girls are more self-efficacious and higher performers than their male counterparts; (see Pajares, Miller & Johnson 1999, and, Britner & Pajares, 2001).

Nevertheless, there is one methodological implication that should be brought to the fore. The original sample size of the current study comprised 122 students; nevertheless, 11 students failed to achieve the minimum attainment threshold of 30%, and were removed. These outlying students’ grades were recorded as an ‘F’ on the data sheet, and not as a percentage score; meaning their sub-optimal attainments could not be quantified. All eleven outlying students were male.

**Student ability levels**

The ability levels of the students proved to be a highly confounding factor in both sets of results. Prior second-year performances were more predictive than self-efficacy measures among weaker performing students (according to t-tests and Pearson’s Coefficient), while self-efficacy proved to be moderately accurate at mapping out end-of-third-year grades for mid-tier students. However, neither variable or statistical calculation drew a positive correlation with higher performing students. This occurrence may be due to the possibility that more assiduous students are aware of the pedagogical implications of the subject matter, set their ambitions well beyond the scope of the current syllabus, or simply may not wish to elevate expectations and fail to meet them. Regardless or the rationale, students improving by greater margins tended to display higher levels of self-efficacy, insinuating that self-efficacy may have been the driving force behind achieving positive change.

**Task-specificity**

Self-efficacy declarations for overall English grades were seemingly accurate, which the data in this study were able to verify. This was also the case on a sub-domain level, as self-efficacy readings closely coincided with scores of both mid-term tests (week 8) and final exams (week 16), which helps to assume that the temporal factor played no intervening role in the results.

However, the most evident incongruity noted among the data were the declared efficacy beliefs with respect to speaking scores, which were considerably understated notwithstanding subsequent high levels of performance. The inconsistencies recorded within this parameter appear to refute Pajares’ (1996) claims hypothesising that task-specific measures offer greater precision. Further study into students’ low confidence levels in specific subject areas warrants further inquiry; namely to investigate the causality of attitudes, perception, environmental and
affective factors in relation to English language attainment in general.

Conclusion and recommendations

The findings in this study discovered that self-efficacy scales are generally accurate when assessing a larger number of students’ beliefs (of varied ability levels) in their overall attainment potential. In more precise terms however, this study helped to expose the extent to which students’ ability levels and task specificity interfere with the accuracy of their respective efficacy judgments. Tschannen-Moran and Hoy’s (2001) assessment summarising that researchers should avoid measures that are highly general, (such as a GPA), or too specific, (such as a speaking task focusing on a specific topic), adequately sums up the mixed results brought to the fore in the current study. Nevertheless, this does not undermine the utility of task-specific self-efficacy scales in the broader context of education. To date, there have been few viable alternative psychological frameworks that better predict or gauge motivational tendencies and the direction of individual behaviour. The fact remains that the relative accuracy and quick deployment of self-efficacy scales offer educators an indispensable tool to ascertain students’ subject-level capabilities and to identify low-confidence areas in the curriculum, helping to better utilise academic resource and guide methodology.

To enhance future findings, the current study was able to identify several confounding factors worthy of further investigation, such as participant profile characteristics, the gender variable, ability levels and task specificity. The main limitation of the current study was that the participants were from a mono-ethnic background studying at one individual higher-education institution in Bangkok, and were from similar socio-economic backgrounds. Four suggestions have been put forth to conduct further-reaching research to reevaluate the strength of self-efficacy scales, and further untangle the confounding elements identified.

Firstly, successive research could be carried out in several different countries in order to ascertain the potential intervention of cultural and socioeconomic factors in self-efficacy development. Secondly, research targeting multiple age levels (elementary, high school, university and adults) may help to observe the evolutionary manner of self-efficacy beliefs as a function of age.

For further insight, the extent to which the evolutionary development of efficacy beliefs differs according to subject matter, such as mathematics, science, and other foreign languages would be worthy of empirical focus. Lastly, added scrutiny could be paid to the role of the gender variable, especially seeing that the inclusion of the outlying data in the present study would have revealed differing results; thus, cross-sectional analyses across a wider demographically balanced sample may elucidate behavioural differences according to gender more clearly.

References


Psychology and Aging, 2(3), 266-271.


Appendix

Efficacy questionnaire administered in the local Thai language.

1. ฉันเชื่อว่าเกรดเฉลี่ยวิชาภาษาอังกฤษของฉันในเทอมนี้คือ …
   (This term, I believe I can achieve an overall English score of …)

<table>
<thead>
<tr>
<th>F</th>
<th>F</th>
<th>F</th>
<th>D</th>
<th>D+</th>
<th>C</th>
<th>C+</th>
<th>B</th>
<th>B+</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td>70%</td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
</tbody>
</table>

2. ในการสอบกลางภาค ฉันเชื่อว่าฉันจะทำคะแนนได้ …
   (In the mid-term exam, I believe I can achieve a score of …)

<table>
<thead>
<tr>
<th>F</th>
<th>F</th>
<th>F</th>
<th>D</th>
<th>D+</th>
<th>C</th>
<th>C+</th>
<th>B</th>
<th>B+</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td>70%</td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
</tbody>
</table>

3. ในการสอบปลายภาค ฉันเชื่อว่าฉันจะทำคะแนนได้ ……
   (In the final exam, I believe I can achieve a score of …)

<table>
<thead>
<tr>
<th>F</th>
<th>F</th>
<th>F</th>
<th>D</th>
<th>D+</th>
<th>C</th>
<th>C+</th>
<th>B</th>
<th>B+</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td>70%</td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
</tbody>
</table>

4. ในการทดสอบการพูด ฉันเชื่อว่าฉันจะทำคะแนนได้ …
   (In the speaking test, I believe I can achieve a score of …)

<table>
<thead>
<tr>
<th>F</th>
<th>F</th>
<th>F</th>
<th>D</th>
<th>D+</th>
<th>C</th>
<th>C+</th>
<th>B</th>
<th>B+</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td>70%</td>
<td>80%</td>
<td>90%</td>
<td>100%</td>
</tr>
</tbody>
</table>