Effects of Different Evaluative Feedback on Students' Self-efficacy in Learning

Joanne C.Y.Chan and Shui-fong Lam

The University of Hong Kong

Submitted on June 19, 2008

Author Note

Effects of Different Evaluative Feedback on
Students' Self-efficacy in Learning

Abstract

Two studies examined the effects of four types of teachers' evaluative feedback on Chinese students' self-efficacy in English vocabulary acquisition. In Study 1, a random sample of Grade 8 students (N=79) learned prefixes and received either formative or summative feedback after failure in test. The results showed that students who received summative feedback showed a larger decrease in their self-efficacy than those who received formative feedback. In Study 2, a random sample of Grade 7 students (N=77) went through similar procedures as in Study 1 except that students received either self-referenced or norm-referenced feedback. The results showed that self-referenced feedback was more beneficial to students' self-efficacy than norm-referenced feedback. The influences of teachers' evaluation and feedback on students' self-efficacy are discussed.

Effects of Different Evaluative Feedback on
Students' Self-efficacy in Learning

In social cognitive theory, personal factors, behavior, and external environment form a triadic reciprocal causation (Bandura, 1986). Self-efficacy theory is concerned with the factor of personal causation within this interdependent structure and provides guidance for individuals to exercise influence over their way of living (Bandura, 1997). By definition, self-efficacy is the belief in an individual's capabilities to execute behavior that is required to achieve prospective outcomes (Bandura, 1977). Self-efficacy theories distinguish among the agents of control, the means of control, and the ends of control (Skinner, 1996). Outcome expectations are individuals' estimates that a certain behavior (means) can achieve the desired outcomes (ends) whereas efficacy expectations are individuals' belief of whether they (agent) can produce the behavior (means) which can produce the desired outcomes (ends) (Bandura, 1977). According to Bandura (1997), though certain behaviors will attain desired outcomes, individuals will not exert control unless they believe that they can produce these behaviors. In this research, we endeavor to test empirically how teachers can manipulate the ends and means of control to affect students' self-efficacy.

Self-efficacy affects one's level of motivation, affective states, actions, thought patterns and resilience (Bandura, 1997; Pajares, 1996). Individuals with high self-efficacy invest more effort and persist longer than those with low self-efficacy, especially when they face setbacks (Bandura, 1989). Individuals with high self-efficacy are also more likely to attribute failure to insufficient effort or situational factors, whereas those with low self-efficacy tend to attribute failure to lack of ability (Bandura, 1999). Given the

powerful influence of self-efficacy, researchers and educators are eager to understand what contributes to its development and maintenance. It is important to note that self-efficacy beliefs are multi-dimensional rather than uni-dimensional (Zimmerman & Kitsantas, 1997)..The present research focuses on the domain of academic self-efficacy.

Academic Self-efficacy and Sources of Efficacy Information

Perceived academic self-efficacy is defined as the personal judgment of one's capabilities to organize and execute the courses of action to attain specific types of educational performances (Zimmerman, 1995). Research has shown that self-efficacy is related positively to motivation, memory, persistence, stress management, cognitive engagement, use of self-regulatory strategies and achievement (Linnenbrink & Pintrich, 2002; Zimmerman, 1995).

Bandura (1997) postulated four fundamental sources of information that consolidate self-efficacy, namely mastery experience, social persuasion, vicarious experience, and physiological and affective states with mastery experience being the most influential. In general, success increases self-efficacy and failure lowers it (Bandura, 1999). As self-efficacy judgments and performance are reciprocally linked, previous performance can alter one's efficacy judgments and one's efficacy can affect subsequent performance (Cervone, 1993).

Besides performance, social persuasion influences individuals' self-efficacy. Social persuasion that confirms the potential of individuals strengthens their self-efficacy but an individual's subsequent poor performance may decrease it (Schunk, 2003). Furthermore, social persuasion that convinces individuals of their lack of ability also leads them to avoid challenge and give up easily (Bandura, 1999). In an academic setting,

teachers and peers are the immediate sources of social persuasion for students. In the current research, we look at how teachers affect students' self-efficacy through manipulation of mastery experience and social persuasion.

Effects of Teachers' Evaluative Feedback on Students' Self-Efficacy

Teachers partially determine students' mastery experience by the evaluative standards they set. Consider a real case scenario in Hong Kong where two schools adopt different passing standards for a class test: one school may set the passing score at 60% and another school may set it at 50%. Thus, students with the same absolute score of 55% would obtain a pass in the former but a fail in the latter. Different reference standards yield different interpretations of success that subsequently influence students' evaluation of their self-efficacy. In addition, teachers' evaluative feedback constitutes social persuasion that affects students' self-efficacy. A teachers who confirms students' capabilities may enhance their self-efficacy but a teachers who tells students that they are incapable may weaken their self-efficacy.

From the perspective of social cognitive theory, teachers' feedback constitutes an environmental variable that influences self-efficacy, a personal variable (Schunk & Zimmerman, 1997; Schunk, 2003). Individuals react to feedback by maintaining successful strategies and modifying unsuccessful ones (Zimmerman, 1997).Research has shown that feedback that informed students of their mastery of learning strategies (Zimmerman & Kitsantas, 2002) and those that linked students' success with their mastery of strategies both contributed positively to self-efficacy (Schunk & Rice, 1987, 1992, 1993). Schunk and colleagues conducted a series of experiments to examine the effects of feedback and goals on students' self-efficacy and found that progress feedback

together with learning goal are beneficial to their self-efficacy (Schunk & Rice, 1991; Schunk & Swartz, 1993a, b). There is also evidence that the sequence of ability and effort attribution feedback affected self-efficacy (Schunk & Rice, 1986).

Given the evidence that feedback affects self-efficacy, we endeavored to identify how feedback affects self-efficacy through Ellen Skinner's (1996) construct of control. Adopting Skinner's terminologies, in an academic setting, students are the agents of control, learning strategies are the means of control and designated educational goals are the ends of control. Feedback that convinces students that they, as the agent, can exercise control over the means enhances efficacy expectation. Moreover, feedback that provides students with means to achieve their goals, or the ends, strengthens outcome expectations. The goal of this paper is to provide empirical evidence of how the triadic construct of control can further our understanding of the effects of teachers' feedback on students' self-efficacy.

(Insert Figure 1 about here)

Importance of Achievement Goals and Feedback on Self-Efficacy

Teachers' evaluative feedback can influence students' self-efficacy through its impact on students' achievement goals. Achievement goals are the ends in the construct presented in Figure 1. According to Ames and Ames (1984), different structures of the learning environment can make different goals salient. Although researchers use different terminologies (e.g. Ames & Archer, 1987, 1988; Dykeman, 1994; Nicholls, 1979; Schunk & Rice, 1989; Schunk & Swartz, 1993a), achievement goals have been generally dichotomized into learning versus performance goals (Dweck, 1986; Dweck & Leggett, 1988; Elliott & Dweck, 1988).

Learning goals define competence in absolute terms (Elliot & McGregor, 2001). Students who adopt learning goals are oriented towards learning and improvement based on a self-referenced standard. They tend to attribute outcome to effort (Ames, 1992). Since students can control how much effort they invest in a given task, they will have more confidence over the agent-means relation, that is, higher self-efficacy. Learning goals can be further differentiated into mastery –approach goal that target at approaching success and mastery-avoidance goal that targets at avoiding failure (Elliot & McGregor, 2001).

Performance goals define competence in normative terms. Students with performance goal usually focus on using ability to achieve success and impress others. Since most individuals believe that ability is relatively stable and unalterable (Weiner, 1986), they may perceive less control over the means of achievement, hence lower their self-efficacy. There is the differentiation between performance-approach goal that aims at approaching success and performance-avoidance goal that aim at avoiding failure (Elliot & Harackiewicz, 1996; Skaalvik, 1997; Elliot & McGregor, 2001). Students may also adopt a failure-avoiding pattern of motivation when they are unsure if others will evaluate their performance positively (Ames, 1992).

Some researchers endorse a multiple goal perspective which claims that learning and performance goals coexist in an individual's goal system (Barron & Harackiewicz, 2001). Urdan and Maehr (1995) also proposed the concept of social goals to describe the social purposes underlying academic achievement. For the purpose of this paper, we focused on mastery-approach goal and performance-avoidance goal. These two goals were focused because they have the most opposite consequences after setback. Mastery

goal is resilient to failure and still elicits mastery behaviors and approach mentality after negative feedback. In contrast, performance goal in face of failure will turn into performance-avoidance goal that elicits helpless behaviors and avoidance mentality. For the parsimony and convenience of presentation, we will refer to the former as learning goal and the latter as performance goal throughout the paper.

Four Types of Evaluative Feedback and the Two Achievement Goals

Achievement goals and evaluative feedback influence each other. Given that each achievement goal creates its own framework for processing information (Elliott & Dweck, 1988), students with different achievement goals may interpret feedback differently. Evaluative feedback, on the other hand, orients students towards different achievement goals. In the present research, we will examine four types of feedback: formative, summative, self-referenced and norm-referenced.

Formative feedback and summative feedback act on the means of control (see Figure 1). Formative feedback provides students with the means to achieve a goal and thus orients students towards a learning goal (McAlpine, 2004). As a result, students may invest more effort or adopt different strategies when they face failure (Elliott & Dweck, 1988). On the contrary, summative feedback focuses on the outcome (McAlpine, 2004) without empowering students with the means to achieve the goal. This focus facilitates a performance orientation and students are more likely to associate their performance with ability evaluation (Ames, 1992) and attribute failure to low ability. According to Schunk and Zimmerman (2007), students' conceptions of ability may influence self-efficacy. When students believe that ability is not malleable (Elliott & Dweck, 1988), they

perceive little control over ability as the means of goal achievement and lower their self-efficacy.

Besides directly acting on the means of control, evaluative feedback can also affect the ends of control, which inevitably influences the means. Classroom structures affect students' perceptions of cognitive abilities mainly by emphasizing social-comparative versus self-comparative appraisal (Bandura, 1997). Schunk (1991) stated that self-referenced comparison helps students concentrate on their progress and adopt learning goal. Since students' improvement in the mastery of specific skills or areas of knowledge is being assessed, they are more likely to focus on using effort to make improvement. Therefore, this sense of control may enhance self-efficacy.

On the contrary, norm-referenced evaluation urges students to demonstrate their ability by outperforming others (Popham, 2001). It leads students to focus on comparative performances and encourages outcome attribution to ability (Ames, 1992). However, as students lack the means of controlling the performance of others, their self-efficacy weakens. After failure, students may adopt performance-avoidance goal to avoid social comparison that may reveal their incapability.

Past research has shown that students in self-referenced groups demonstrated higher self-efficacy and better skill than those in social-referenced groups (Shih & Alexander, 2000). It was reasoned that the self-referenced groups could concentrate on self-improvement and engaged actively in the task but the social-referenced groups were preoccupied with comparing their abilities with others, resulting in anxiety which interfered with both cognitive engagement and self-efficacy.

Contributions of the Current Research

As Bandura (1986) pointed out, one of the objectives in academic research is to study how different types of educational practices and structures affect the development of social and cognitive competencies. Over time, it is believed that students who have developed a strong sense of self-efficacy are better equipped to progress with their own initiatives. Therefore, it is of both theoretical and practical importance to study self-efficacy in academic settings.

Against this backdrop, this research aimed at responding to the call for a better understanding of adolescents' self-efficacy in specific settings, which has been underrepresented in the literature (Anderman & Maehr, 1994; Pajares, 1996). We looked into the self-efficacy of Chinese adolescents in English vocabulary acquisition. With reference to the triadic relation among the agents, means, and ends in control, we investigated the effects of teacher's evaluative feedback on students' self-efficacy.

To bring about a greater capacity to generalize the findings, we chose four types of evaluation feedback that are commonly used in classrooms, namely summative feedback, formative feedback, norm-referenced feedback, and self-referenced feedback. We also used normal students as participants to complement the substantial research that looked at students with special needs (e.g. García & de Caso, 2004; Graham, Harris, & Mason, 2005; Schunk, 1982; Schunk & Swartz, 1993b; Schunk & Rice, 1986, 1987, 1989, 1993).

We chose vocabulary acquisition as the subject of investigation because it is an essential element in language acquisition (Penno, Wilkinson, & Moore, 2002) and vocabulary knowledge is a well-established predictor of reading comprehension (e.g.

Cunningham, & Stanovich, 1997; Spearitt, 1972) as well as academic progress (Walker, Greenwood, Hart, & Carta, 1994).

The present research employed experimental controls in a field setting with Chinese-Confucian Heritage which has a strong emphasis on effort and endurance (Hau & Salili, 1996). Classrooms with this heritage are also known for their competitiveness and high prevalence for normative evaluation (Biggs, 1996; Shih & Alexander, 2000). Therefore, this research would shed light on how different types of feedback can affect students' self-efficacy in a competitive learning environment.

The two studies in this research were conducted in actual classrooms to advance their internal and external validities. The findings strive to enable instructors to design and implement evaluative feedback that can improve students' self-efficacy and consequently enhance their learning motivation and achievements.

Overview of the Two Studies

The present research investigated the effects of four types of evaluative feedback on students' self-efficacy. The first study focused on the manipulation of means in the concept of control. Using formative and summative feedback, we compared the effects of providing and withholding means of control on students' self-efficacy. The second study focused on the manipulation of ends in the concept of control. Using self-referenced and norm-referenced feedback, we compared the effects of two different ends, namely outperforming self and outperforming others, on students' self-efficacy.

Study 1

In Study 1, we compared the effects of summative and formative feedback on students' self-efficacy. Formative feedback that provides strategies for subsequent improvement gives students the means to achieve a goal explicitly while summative feedback that addresses past performance that is no longer controllable withholds such means from students. As self-efficacy is concerned with an agent's perceived control over the means towards a designated goal, we hypothesized that the self-efficacy of students in the formative feedback condition would be higher than that in the summative feedback condition.

Method

Participants

Participants were 79 Chinese students in Grade 8 who joined a vocabulary builder program. They were recruited from a secondary school in Hong Kong that uses English as the medium of instruction. Over a quarter of the secondary schools in Hong Kong use English as the medium of instruction so the sample is considered to be representative of these schools. Invitation letters were sent to the parents of all students in the eight classes of Grade 8. Students were assigned to these eight classes based on their English proficiency. Parents were informed that their children might be selected randomly to participate in a "Vocabulary Builder Program" during which they would learn vocabulary acquisition strategies. They were also told that the researchers would study what instructional strategies would be most effective. All parents gave their consent for their children's participation. The English teachers of the eight classes were asked to draw from a hat the names of five male and five female students from their respective class. As

a result, 80 students were recruited but one student was absent on the day of the experiment so only 79 students participated.

Experimental Design

The 79 students were assigned randomly to either the summative feedback condition (n=39) or the formative feedback condition (n=40).The instructor and sequence of instruction were the same for both conditions. The only difference was that one condition received summative feedback and the other received formative feedback in the evaluation of the tests in the program.

Vocabulary Acquisition Program

The program was held in the school computer room during an English lesson. Two assistants, unaware of the experimental procedures, distributed the questionnaires. Upon arrival at the classroom, the instructor assigned each student a desktop computer in a carrel with partitions to separate students. Each computer was installed with a program designed for the experiment. The program presented students with instruction, exercises, and tests of prefix usage. The software automatically recorded the students' test answers.

Procedures

During the program, the students went through two 20-minute instructional sessions that included exercises and tests. After each session they answered a questionnaire. In the first session the instructor introduced six prefixes with examples to illustrate their usage. For example, the prefix "mis" means "bad" or "wrong"; therefore, "misbehave" means behave badly and "miscount" means count wrongly. The examples were in a multiple-choice format (e.g., " I *misplace* my books", the word "misplace"

means a) place correctly, b) put in the wrong place, c) read, d) lost). Students tried to answer them and the instructor later revealed the correct answers.

After teaching the first three prefixes, the instructor taught students how to differentiate between real and pseudo prefixes, that is, words that start with the spelling of a prefix. Then the instructor continued to teach the remaining three prefixes. After that, the instructor taught another strategy. After the teaching session, students responded to the first questionnaire (Questionnaire A) that tapped their self-efficacy with regard to the subsequent teaching session and test. Then students did Test 1 and had 30 seconds to answer each of the 16 questions.

At the end of Test 1, the computers provided students with feedback on their performance on Test 1. Students in the summative condition received summative feedback whereas students in the formative condition received formative feedback. After students had viewed the results, the instructor started the second instructional session and taught another six prefixes and two strategies. Prior to Test 2, the instructor asked students to respond to the second questionnaire (Questionnaire B). The questions tapped their sense of mastery with regard to the result of Test 1 and their self-efficacy about the subsequent teaching session and test. Manipulation check items were also included. The students then did Test 2, which had the same format as Test 1. After that, the computer generated genuine feedback based on students' performance. The instructor then informed students that there was not enough time for further instruction. At the end of the school day, we debriefed students from both conditions and informed them that, for the purpose of the study, every student was looking at identical test results for Test 1 that did not reflect their performance accurately. We gave students a package that included all the

prefix teaching material, test items, and answer key for them to assess their actual level of mastery as well as for future reference.

Feedback Manipulation

After Test 1, the instructor told the students in the summative condition that the computers would show them the results, including the number of correct and incorrect answers, the overall percentage of accuracy, and a list of the correct and incorrect items. She also reminded students that it was most important to do well in the test as it could indicate their ability. The results presented by the computers were pre-programmed. All students in the summative condition received identical feedback irrespective of their actual performance. To students in the formative condition, the instructor said that the computers would show them the results, including the number of correct and incorrect answers and some suggestions that were allegedly tailor-made for the students so that they could make improvements. She also reminded students that it was most important to learn from the suggestions as the test could help them find ways of improvement. The students were led to believe that the computers had analyzed their work and made suggestions for their improvement. For example, participants would see the following feedback, "For questions 4, 9, 12, you should try to use the strategy of understanding the meaning of the whole sentence before choosing the answer." These suggestions were identical for all students in the formative condition irrespective of their actual performance. We manipulated the feedback so that the students in both conditions got seven correct answers out of 16 questions in Test 1. According to the grading system in most schools in Hong Kong, such a result would be regarded as a failure because the accuracy rate was below 50%.

Measures

Self-evaluation of success. As mastery experience is the most important source of efficacy information (Bandura, 1997), we asked students to evaluate how the results of Test 1 made them feel in Questionnaire B. The ratings were made on a seven-point scale from 1 for "not successful at all" to 7 for "very successful."

Self-efficacy. Students evaluated their self-efficacy twice subsequent to learning sessions and tests. In the first questionnaire (Questionnaire A) they evaluated their confidence in and control over learning well in the first learning session and getting good results in Test 1. In the second questionnaire (Questionnaire B) they responded to the same questions with regard to the second learning session and Test 2. Specifically, students responded to the following four questions: 1) How confident are you to do well in the next test? 2) How much control do you have over the result of the next test? 3) How confident are you to learn all the prefixes in the next lesson? 4) How much control do you have over how well you learn in the next lesson? Seven-point scales from 1 for "not confident at all" or "no control at all" to 7 for "very confident" or "very much control" were used. The Cronbach's alpha of these four items was .88 before Test 1 and .89 before Test 2; showing a high internal reliability.

Manipulation checks. In Questionnaire B, we asked students to assess the quality of instruction and classroom management to see if there were any differences between the two conditions. The evaluation of the clarity of teaching was on a 7-point Likert scale from 1 for "not clear at all" to 7 for "very clear." For noise level of the classroom, students rated from 1 for "not noisy at all" to 7 for "very noisy." We also asked students to indicate if they had learned more prefixes after the two lessons to evaluate the

effectiveness of our instruction. Moreover, we asked them to indicate the number of correct answers they had in Test 1.

Results

Performance

We compared students' actual performance in Test 1. The summative condition ($\overline{X}$ = 6.55, SD = 1.72) and the formative condition ($\overline{X}$ = 5.82, SD = 1.57) did not show a significant difference, $t(75)$ = -1.95, p = .06, ns, Cohen's d = .45. In Test 2, the performance of the students in the formative condition ($\overline{X}$ = 11.05, SD = 1.45) were also similar to that of the students in the summative condition ($\overline{X}$ = 11.03, SD = 1.73), $t(75)$= .07, p = .95, ns, Cohen's d =.01.

Self-evaluation of success

After Test 1, students in the formative condition ($\overline{X}$ = 3.59, SD = .72) felt more successful than their counterparts in the summative condition ($\overline{X}$ = 3.15, SD = 1.15), $t(77)$ = 2.04, p = .045, Cohen's d = .46. The difference was statistically significant and the effect size was medium according to statistical conventions (Urdan, 2001).

Self-efficacy

We compared students' self-efficacy before Test 1 and Test 2 (see Table 1). We performed a mixed model Univariate Analysis of Variance (ANOVA) with self-efficacy (Test 1 vs. Test 2) as the within-participant factor and condition (formative feedback vs. summative feedback) as the between-participant factor. The ANOVA yielded a statistically significant main effect for self-efficacy, $F(1, 77)$ = 49.41, p < .001, partial η^2 = .39, but not for condition, $F(1, 77)$ = .001, p > .05, η^2 < .01. However, the interaction

between self-efficacy and condition was at the .05 statistical level, $F(1, 77) = 3.92$, p $= .05$, $\eta^2 = .05$.

To understand this two-way interaction, we compared students' self-efficacy between Test 1 and Test 2 (see Table 1). Participants in both conditions experienced a decrease in self-efficacy from Test 1 to Test 2. The formative condition decreased from 5.22 ($SD = .80$) to 4.85 ($SD = .72$), $t(38) = 4.34$, $p < .01$, Cohen's d $= .50$. The effect size was medium. The summative condition also decreased from 5.37 ($SD = .87$) to 4.69 ($SD = .80$), $t(39) = 5.57$, $p < .01$, Cohen's d $= .81$. It is noteworthy that the effect size was large. The summative condition showed a larger decrease in self-efficacy ($\overline{X} = .68$, $SD = .77$) than the formative feedback condition ($\overline{X} = .38$, $SD = .54$), $t(77) = -1.98$, $p = .05$, Cohen's d $= .45$. The difference was significant and the effect size was medium.

(Insert Table 1 about here)

Manipulation Checks

The summative condition ($\overline{X} = 5.75$, $SD = 1.10$) and the formative condition ($\overline{X} = 5.97$, $SD = .78$) both rated the teaching as quite clear, $t(70.2) = 1.04$, $p = .30$, *ns*, Cohen's d $= .23$. Both the summative condition ($\overline{X} = 1.90$, $SD = 1.15$) and the formative condition ($\overline{X} = 1.56$, $SD = .96$) also rated the noise level of the classroom as very low $t(77) = -1.39$, $p = .17$, *ns*, Cohen's d $= .32$. In the formative condition, 92.3% of the students indicated that they had learned more prefixes after the two lessons while 90% of the students in the summative condition claimed so, $\chi^2(79) = .13$, $p = .72$, *ns*. When asked to indicate their results on Test 1, 94% of the formative condition and 94.6% of the summative condition accurately recalled that they got seven correct answers, $\chi^2(76)$

= .003, p = .96, *ns*. This showed that the majority of students remembered that they got less than half of the questions correct in Test 1.

Discussion

The results showed that though students in the formative and summative conditions did not show significant differences in their self-efficacy before Test 1 and Test 2, their change in self-efficacy was significantly different with a medium effect size of .45. As seen, notwithstanding that both conditions had identical learning environments and poor test performances, different evaluative feedback had different effects on students' self-efficacy and sense of success.

In the formative condition, the feedback presented students with concrete means for improvement that strengthened the means-ends relation in the construct of control. Moreover, by telling students that the suggestions for improvement were tailor-made for them, the agent-means relation could also be reinforced. It was possible that, as a result, the consolidation of the triadic relation among the agents, means and ends of control alleviated the decrease in students' self-efficacy after failure. In the summative condition, feedback was presented to students subsequent to their performance without any suggestion for improvement. Since nothing could be done to alter past performance, the means-ends relation would be weak. Students as the agents also had no control over the means to alter the quality of the end product, thus the agent-means relation would also be weak. As a result, students' self-efficacy decreased.

We assumed that students interpreted the result of having only seven correct answers out of 16 questions as failure because it was below the normal passing standard of 50% in most secondary schools in Hong Kong. It was possible that some students

might be attempting to predict their own success and did not see the pre-programmed results as credible. Nevertheless, the actual performance of the students in both summative condition ($\overline{X} = 6.55$, $SD = 1.72$) and formative condition ($\overline{X} = 5.82$, $SD = 1.57$) were in fact even lower than the pre-programmed feedback. Therefore, the feedback would have boosted their self-efficacy instead of dampening it. To improve the accountability of the feedback manipulation, we could have asked students to indicate the credibility of the feedback. To address this limitation, we included this manipulation check in Study 2.

To rule out the possibility that students' self-evaluation and self-efficacy were affected by their actual performance, we compared their actual performance on both Test 1 and Test 2. Students in the two conditions had similar performances on both tests. Therefore, different feedback did not affect actual performance. Nevertheless, students in the formative condition showed a higher sense of success than those in the summative condition. Their self-efficacy also did not decrease as much as those of their counterparts in the summative condition. Different types of evaluative feedback might have had different effects on the mastery experience of students which consequently affected self-efficacy.

In Study 1 we provided the means of control to some students by giving them formative feedback and withheld it from others by providing summative feedback. To further explore further the effects of the manipulation of ends of control on students' self-efficacy, we investigated another two types of feedback in Study 2.

Study 2

In Study 2, we used different feedback to highlight different achievement goals as the ends in the construct of control. We used self-referenced feedback to highlight learning goal and norm-referenced feedback to highlight performance goal. Oriented towards a learning goal, students aim to develop their own competence and are more inclined to make effort attribution. As effort is a controllable factor and students can exercise effort to attain their goal, their self-efficacy can be strengthened. The performance goal, however, exerts pressure on students to outperform others and induces ability attribution. As students lack the means to control their own ability as well as the performance of others, their self-efficacy would be at risk. Hence, we hypothesized that the self-efficacy of students in the norm-referenced feedback condition would be lower than that in the self-referenced feedback condition.

Method

Participants

Participants were 77 Chinese students in Grade 7 who joined a vocabulary builder program. They were recruited using the same procedures from the same school as in Study 1. All parents gave their consent for their children's participation. As a result, 80 students were recruited. However, only 77 students showed up on the day of the experiment.

Experimental Design

The 77 students were randomly assigned to either the self-referenced feedback condition (n=38) or the norm-referenced feedback condition (n=39). The instructor and sequence of instruction were the same for both conditions. The only difference was the evaluative feedback for the tests. The self-referenced condition received feedback in

terms of the number of correct answers and the percentage of accuracy. In contrast, the norm-referenced condition received feedback in terms of the number of correct answers and the percentile ranking when compared to other students.

Vocabulary Acquisition Program

Except where otherwise stated, we used the same material for instruction, test format, questionnaires, and procedures as in Study 1.

Procedures

The experiment took place during a normal English lesson in the school computer room. First, students did a baseline vocabulary test for subsequent evaluation of whether the initial performances of the two groups were comparable. Then the students went through two 20-minute instructional sessions; each included a set of exercises and a test. After completing each test, students answered an online questionnaire. Since the computers malfunctioned in our trial run before the experiment, we also asked the students to record their answers in pencil on the hard copies of the questionnaires. In the first session, the instructor presented the meaning of six prefixes and examples to illustrate their usage.

After learning six prefixes, students did Test 1. After they completed Test 1, the computers gave feedback about their performance. After viewing the feedback students responded to the first questionnaire (Questionnaire A) both online and on paper. The questions tapped their sense of success about their previous performance and their self-efficacy with regard to the subsequent instructional session and test (i.e. the second instructional session and Test 2).

After students had completed the questionnaire, the instructor started the second session to introduce another six prefixes. Then students did Test 2. The format of Test 2 was the same as the previous two tests. At the end of Test 2, the computers provided the two conditions with different feedback. Then students completed the second questionnaire (Questionnaire B) both online and on paper. It tapped their sense of success about their previous performance and their self-efficacy with regard to the subsequent instructional session and test (i.e. the third instructional session and Test 3). After collecting the questionnaires, the instructor informed students that there was not enough time for the subsequent instructional session and Test 3. The instructor then debriefed both groups as in Study 1 and told the norm-referenced condition that comparison with others was actually less important than the mastery of new skills.

Feedback Manipulation

At the end of Test 1, the instructor reminded the students in the norm-referenced condition that their relative performance in comparison with other students was most important. The instructor reminded students in the self-referenced condition that making improvement in comparison with their previous performance was most important. The computers then displayed the number of correct answers for students. We manipulated the feedback so that students in both conditions had six questions correct in the baseline test and ten questions correct in Test 1. In addition, the computers showed the percentage of accuracy for the self-referenced condition and the percentile ranking of the individual's performance when compared to other students for the norm-referenced condition. Moreover, the computers displayed a bar chart that compared the results between the baseline test and Test 1. For the self-referenced condition, the bar chart

showed 38% accuracy in the baseline test vs. 63% accuracy in Test 1. For the norm-referenced condition, the bar chart showed 38th percentile ranking in the baseline test vs. 63rd in Test 1.

At the end of Test 2, the computers showed the results for both groups in a similar format as Test 1. In addition, the computers displayed a bar chart of the comparison among the results of the baseline test, Test 1, and Test 2. We manipulated the feedback so that students in both conditions had 12 correct answers in Test 2. For the self-referenced condition, the bar chart showed the percentage of accuracy among the three tests which were 38%, 63%, and 75% respectively. For the norm-referenced condition, the bar chart showed the percentile ranking among the three tests which were 38th, 63rd and 35th respectively. The fact that an increase in the absolute number of correct answers corresponded to a drop in ranking in Test 1 but an increase in Test 2 was a reflection of the fact that percentile ranking can fluctuate despite improvement in actual scores.

Measures

Self-evaluation of success. In both questionnaires students were asked how the results of the previous test (i.e. Questionnaire A asked about Test 1 and Questionnaire B asked about Test 2) made them feel. The ratings were made on a seven-point scale from 1 for "not successful at all" to 7 for "very successful."

Self-efficacy. Students rated their self-efficacy with regard to the subsequent lessons and tests (i.e. after Test 1, they evaluated their self-efficacy with regard to Test 2 and the learning before that test; after Test 2, they rated that for Test 3 and the learning before that test). We used the same four questions as in Study 1. The Cronbach's alpha of

these four questions was .79 before Test 2 and .85 before Test 3, demonstrating a high internal consistency.

Manipulation checks. In both questionnaires, we asked students whether the previous test result (i.e. Questionnaire A asked about Test 1 and Questionnaire B asked about Test 2) could show how much they had learned about prefixes. This provided us insight into whether students perceived the artificial feedback to be credible. The ratings were made on a seven-point scale, from 1 for "not at all" to 7 for "very much." In Questionnaire B, students indicated whether they had learned more prefixes after the two lessons and the number of correct answers they got for Test 2. The purpose was to see if they recalled that they had the majority of the questions correct. Students also evaluated the clarity of teaching and noise level of the classroom on a seven-point scale as in Study 1.

Results

Performance

We compared students' actual performances on the two tests. Unfortunately, some computers had technical problems and failed to record the test results of about one-third of the students. In the self-referenced condition, the results of 25 students were intact and 13 were lost. In the norm-referenced condition, the results of 26 students were intact and 13 were lost. To ensure that the group whose test results were lost was comparable to the group whose test results were retained, we examined the manipulation checks of these two groups of students. Within each condition, these two groups of students had no difference in their perception of the credibility of their test results, the clarity of the instruction, the noise level of the classroom, how much prefixes they had learned, and the

number of correct answers they got in Test 2, all $p > .05$. The analyses indicated that the attrition in test results had not produced a biased sample.

In the baseline test, the performance of the self-referenced condition ($\overline{X} = 8.28$, $SD = 3.29$) and the norm-referenced condition ($\overline{X} = 8.33$, $SD = 2.47$) were comparable, $t(50) = -.067$, $p = .95$, ns, Cohen's d = .02. In Test 1, the performance of the self-referenced condition ($\overline{X} = 8.20$, $SD = 1.98$) and the norm-referenced condition ($\overline{X} = 9.00$, $SD = 1.14$) were also comparable, $t(50) = -1.80$, $p = .078$, ns, Cohen's d = .51. In Test 2, the performance of the students in the self-referenced condition ($\overline{X} = 9.20$, $SD = 2.25$) and the norm-referenced condition ($\overline{X} = 8.67$, $SD = 2.08$) also did not differ, $t(50) = .89$, $p = .38$, ns, Cohen's d = .31.

Self-evaluation of success

Besides performance, the rest of the data was intact. We compared students' self-evaluation of their success after Test 1 and Test 2. The self-referenced condition ($\overline{X} = .43$, $SD = 1.27$) felt more successful than the norm-referenced condition ($\overline{X} = -.24$, $SD = 1.36$) and the difference was statistically significant with a medium effect, $t(70) = 2.16$, $p = .034$, Cohen's d = .52.

Self-efficacy

We performed a mixed model Univariate Analysis of Variance (ANOVA) with self-efficacy (Test 2 vs. Test 3) as the within-participant factor and condition (self-referenced vs. norm-referenced feedback) as the between-participant factor. The ANOVA yielded a significant main effect for condition, $F(1, 73) = 9.31$, $p < .01$, partial $\eta^2 = .11$. The main effect for self-efficacy was not significant, $F(1, 73) = 3.75$, $p = .06$, partial η^2

= .05 but the two-way interaction between self-efficacy and condition was significant,

$F(1, 73) = 45.45$, $p = .001$, $\eta^2 = .39$.

To understand this two-way interaction, we decomposed it into a series of t-tests, examining participants' self-efficacy for Test 2 and Test 3 (see Table 2). The self-referenced condition experienced a significant increase in self-efficacy from Test 2 ($\overline{X} = $ 4.83, SD = .84) to Test 3 ($\overline{X} = 5.39$, SD = .79), $t(36) = 6.55$, $p < .001$, Cohen's d =.67. The effect size was medium. In contrast, the norm-referenced condition suffered a significant decrease in self-efficacy from Test 2 ($\overline{X} = 4.72$, SD = .78) to Test 3 ($\overline{X} = 4.41$, SD = .89), $t(37) = -3.22$, $p < .001$, Cohen's d = -.37. The two conditions started out with no difference in self-efficacy for Test 2 (4.83 vs. 4.72), $t(74) = .59$, $p > .05$, Cohen's d = .14. However, their self-efficacy for Test 3 after the feedback showed a significant difference with a large effect size (5.39 vs. 4.41), $t(74) = .59$, $p < .05$, Cohen's d = 1.18. The difference between self-efficacy for Test 2 and Test 3 was .56 (*SD* = .52) for the self-referenced condition but -.32 (*SD* = .59) for the norm-referenced condition. The difference between the two conditions was statistically significant and with a large effect size, $t(74) = 6.89$, $p < .001$, Cohen's d = 1.6.

(Insert Table 2 about here)

Manipulation Checks

Students indicated the extent to which the test results reflected their prefix acquisition. The self-referenced condition rated both tests positively (above the neutral point of four) for showing how much they had learned (for Test 1, $\overline{X} = 4.66$, *SD* = 1.26 and for Test 2, $\overline{X} = 4.97$, *SD* = 1.48) and the same applied to the norm-referenced condition (for Test 1, $\overline{X} = 4.73$, *SD* = 1.28, and for Test 2, $\overline{X} = 4.49$, *SD* = 1.35). There

was no significant difference between the two groups' evaluation for both Test 1, $t(70) =$ -.24, $p = .81$, Cohen's d = .06, and Test 2, $t(74) = 1.49$, $p = .14$, Cohen's d = .34. Thus we inferred that the results of both tests appeared to be credible for both conditions.

There were 92.1% of students in the self-referenced condition and 89.7% of students in the norm-referenced condition who indicated that they had learned more prefixes after the two lessons. The difference between the two conditions was not statistically significant, χ^2 (77) = 1.32, $p=.52$. Furthermore, 94.9% of the self-referenced condition and 94.6 % of the norm-referenced condition remembered accurately that they had 12 answers correct in the previous test and there was no statistically significant difference, χ^2 (76) = .003, $p = .96$.

The two conditions only differed in their evaluation of the clarity of instruction. The norm-referenced condition rated the teaching as above the level of being "quite clear" ($\overline{X}$ = 5.44, SD = 1.12), while the self-referenced condition gave an even higher rating ($\overline{X}$ = 6.08, SD = 1.01) and the difference was statistically significant with a medium effect, $t(74) = 2.63$, $p = .01$, Cohen's d =.61. With regard to the noise level of the classroom, the self-referenced condition ($\overline{X}$ =2.41, SD = 1.21) and the norm-referenced condition ($\overline{X}$ = 2.85, SD = 1.50) showed no significant difference, $t(74) = -1.41$, $p = .16$, Cohen's d = .33.

Discussion

In Study 2, as hypothesized, students who received norm-referenced feedback showed lower self-efficacy than those who received self-referenced feedback. The difference in the change in self-efficacy between the two conditions was significant with a large effect size of 1.6.

In the self-referenced condition, students as the agent of control used effort as the means to reach their learning goal, the ends of control. This belief in the agent-means relation strengthened self-efficacy. On the contrary, students in the norm-referenced condition were constantly being compared to other students. As it impossible for them to perceive control over the performance of others, they had no means to outperform others. Thus, their self-efficacy decreased. Despite the fact that students observed their progressive improvement in terms of the absolute number of correct answers, the fluctuation in their percentile ranking had reinforced their helplessness. This finding echoes the observation of Elliott and Dweck (1988) that when individuals believe that they have low ability, they will react in a learned helplessness manner in face of performance goal.

There is concern that the evaluative feedback might have presented a success scenario for the self-referenced condition but a failure scenario for the norm-referenced condition. However, it is important to note that both groups received identical results in terms of the number of correct answers throughout the program. Nevertheless, the type of evaluative feedback administered did have an effect on the students' self-evaluation of success. The self-referenced condition showed an overall increase in the sense of success while the norm-referenced condition showed an overall decrease, with a medium effect size of .52. This difference may be because the norm-referenced condition perceived no means to control others' performance, and thus their own percentile ranking. Consequently, they were more likely to interpret the experience as a setback.

To rule out the possibility that the self-efficacy and sense of success of students were influenced by their actual performance, we compared their actual performance in

Test 1 and Test 2. Within the two-thirds of data retained, there was no difference between the two conditions.

In the manipulation check, the norm-referenced condition rated the clarity of teaching as above "quite clear," while the self-referenced condition gave an even higher rating. Given that both groups perceived the clarity of teaching as above "quite clear," we had confidence that the instructions were clear enough.

General Discussion

The present paper investigated how different types of evaluative feedback affected students' self-efficacy in vocabulary acquisition. Through different types of evaluative feedback, teachers oriented students towards different achievement goals. Formative and self-referenced feedback encouraged students towards a learning goal that emphasized incremental self-improvement; on the contrary, summative and norm-referenced feedback pointed towards a performance goal that defined success as outperforming others.

The results were consistent with findings in the literature on goal-theory of achievement motivation (Ames, 1992). In Study 1, formative feedback provided strategic information for students to make improvement. Learning strategies, as the means, promised a sense of control over academic goals and strengthened the means-ends relation in control. Thus, students, as agents, would perceive more control in the agent-means relation which could either enhance their self-efficacy or lessen the decrease in self-efficacy when they face setbacks. In contrast, summative feedback prompted students to produce a quality end product without providing strategic information. Students were less likely to perceive control over the means-ends relation, as well as the

agent-means relation. Hence, it would be more difficult for them to maintain their self-efficacy.

These findings imply that students' self-efficacy can benefit from teachers' use of formative feedback. Specifically, instead of summarizing the merits and mishaps of students at the end of their performance, teachers can try to provide strategies for students to make improvement during a task. This may be coupled with other techniques, for example, verbal persuasion like "You can do it" (Schunk & Zimmerman, 2007) to convince students of their capability.

The findings of Study 2 also cast legitimate doubts on the appropriateness of norm-referenced feedback wherein student performance is evaluated comparatively (Marsh, Kong, & Hau, 2000). In reality, teachers with their expertise and power (Oettingen, 1995) control the assessment criteria and mode of evaluation (Buunk, Collins, Taylor, VanYperen, & Dakof, 1990) and students passively receive information about how their capabilities are compared to their peers' (Bandura, 1995). As explained earlier, such information affects students' self-appraisal of ability, and thus their academic self-efficacy (Oettigen, 1995). Therefore, teachers should consider using self-referenced feedback instead.

Limitations and Merits

The present study has demonstrated how formative, summative, norm-referenced, and self-referenced feedback could deliver immediate impact on students' self-efficacy despite the brief evaluative session which might not be representative of an actual school experience. Though this limited timeframe of intervention showed effects of different feedback, it fell short of demonstrating whether these effects might be maintained over

time. Nevertheless, if such feedback were to be administered regularly over time, there is reason to believe that there would be long-term effects on students. The brief experiments, which consisted of only two consecutive class periods, limited any prolonged detrimental effects that the norm-referenced and summative measures might have on the participants.

Students within the same experimental condition received identical and pre-programmed feedback throughout both studies. The only exception was the last test of Study 1 in which students received genuine feedback about their performance. The pre-programmed feedback might lead some students to question the credibility of the results. However, the manipulation checks in Study 2 indicated that students generally found the results credible and reflective of their actual performance.

With reference to previous experimental studies (Butler, 1993; Elliott & Dweck, 1988), the present research highlighted learning and performance goals through verbal instruction and evaluation objectives. Measuring goal orientation was duly considered, but ultimately rejected for the following reasons: First, previous research has shown that even when students come to an experiment with learning goal, experimental manipulation in the form of competition, for example, that activates performance goal might still affect learning motivation and result in the worsening of self-evaluation after failure (Lam, Yim, Law, & Cheung, 2004). This provides evidence that the teacher-assigned goals are likely to override the individual goals that students bring to the classroom. Second, the specificity and transient nature of classroom goals or task goals that were the major foci of the present research made the stable goal orientation inadequate as it might not reflect the students' on-task goals.

In the present studies, the scale that measured self-efficacy had students evaluate their level of confidence in or control over learning and performance. The purpose was to highlight the agent-means relation in the concept of control. In fact, self-efficacy is the agent's perceived controllability over the means of goal achievement. We took into account several criteria in developing this scale. The first criterion was that judgments of self-efficacy are task and domain specific so measures of self-efficacy should be tailored specifically to the task being assessed (Pajares & Miller, 1995). Hence, the questions prompted students to evaluate their self-efficacy with regard to "the next lesson" and "the next test." The aim was to focus their attention on evaluating their self-efficacy with regard to the task alone.

The second criterion was that self-efficacy is about prospective situations (Bandura, 1997). Hence, all of the questions on self-efficacy were written in the future tense to highlight the concern about future performance rather than past or present performance. Students were asked to evaluate their self-efficacy specifically with regard to the learning of the prefixes and doing well in the upcoming lessons and tests. The high internal reliability of the self-efficacy items was reflected in the Cronbach's alpha coefficients ranging from .79 to .89.

In response to the demand for more studies on self-efficacy in actual classrooms (Schunk & Zimmerman, 1997), we endeavored to conduct tightly controlled experiments in a classroom on during school hours to maximize internal and external validities. We chose vocabulary acquisition as the object of learning by virtue of its value in language acquisition. Despite its importance, this area has been neglected while research on academic self-efficacy has examined reading and comprehension (e.g. Schunk & Rice,

1986, 1987, 1989), writing (e.g. Schunk & Swartz, 1993a, 1993b), learning mathematics (e.g. Hackett, 1985; Reyes, 1984; Schunk, 1982; Zimmerman & Martinez-Pons, 1990), and computer skill acquisition (Schunk & Ertmer, 1999). Previous research on students' self-efficacy in literary activities seldom looked at how students acquire a second language. More specifically, this study focused on how Chinese students learned English vocabulary.

Cultural Factors

The present research employed experimental controls in a field setting with Chinese students as participants. Contemporary Chinese students, teachers, and parents are all aware of the competition that prevails on the path to higher levels of schooling (Stevenson & Lee, 1996). Students in this culture are accustomed to the competition and high prevalence of normative evaluation in classrooms (Biggs, 1996; Shih & Alexander, 2000). Nevertheless, the participants were still not immunized against the negative effects of norm-referenced feedback in Study 2. Teachers should endeavor to minimize such negative effects when they design and conduct evaluation.

Besides Chinese culture, other cultures also rely heavily on normative assessment (France-Kaatrude & Smith, 1985). For instance, American middle schools emphasize relative ability and social comparison in learning (Midgley, Anderman, & Hicks, 1995) where the success and failure of students are determined comparatively (Levine, 1983; Richer, 1976). Hence future studies can investigate the effects of teachers' feedback on students' self-efficacy across cultures under similar educational regimes.

Future Research Directions

Schunk and Zimmerman (1997) called for more research on the role of social origins such as models, verbal description, social guidance and feedback in students' academic competence. In response, this paper investigated the effects of teachers' feedback as an important source of students' mastery experience and social persuasion on students' self-efficacy. Future research would entail studies to gain a better understanding of other sources of self-efficacy and the relation among them. For example, researchers have already shown that enhancing the physiological and psychological well-being of students and reducing their negative emotions can strengthen self-efficacy (Usher & Pajares, 2006). It is therefore desirable to obtain more understanding about how students can improve their physical and emotional well-being as well as how to empower students to interpret their physiological and psychological experiences positively.

According to Ames (1992), teachers' own goal orientation can influence their beliefs about the efficacy of instructional decisions and teaching methods. Future research should address the impact of teachers' goal orientation and their preference for evaluative feedback on students' self-efficacy, including advantages and disadvantages of alternative evaluations. As teachers play an important role in students' goal-setting and self-efficacy, more research needs to address the effects of different classroom practices and the interaction between teachers and students.

Adolescents spend a major portion of their time at school, where they are constantly exposed to school-related stress such as academic standing, examinations, competition and conflict among peers, and relationships with teachers (Fanshawe & Burnett, 1999; Thuen & Bru, 2004). As self-efficacy is related positively to positive

schooling outcomes such as motivation and achievement, more research is needed to address the self-efficacy of adolescents at school.

Conclusion

For decades, educators have struggled with the challenge of making the learning environment effective for early adolescent students (Urdan, Midgley, & Wood, 1995). Considering how to enhance students' self-efficacy may be a step towards building an effective learning environment. In this research, we showed that teachers, through different evaluative feedback, affected the self-efficacy of students. Formative and self-referenced feedback that enabled students to perceive a sense of control over their progress was beneficial to their self-efficacy. Conversely, summative and norm-referenced feedback that weakened students' perceived control over the achievement outcome threatened their self-efficacy. According to Bandura (1997), students' appraisals of their capabilities are largely affected by the relative emphasis on social comparative versus self-comparative appraisal they receive in the classroom. "Socially competitive grading practices convert educational experiences into ones where many are doomed to failure for the high success of a few."(Bandura, 1997, p.175). Having the responsibility to design and provide feedback, teachers could employ more self-referenced evaluation and provide more learning strategies in their feedback to empower the self-efficacy of students.

Word-length of the paper without bibliography: 8,862

Word-length of the paper with bibliography: 10,572

References

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Education Psychology, 84,* 261-271.

Ames, C., & Ames, R. (1984). Systems of student and teacher motivation: Toward a qualitative definition. *Journal of Educational Psychology, 76,* 535-556.

Ames, C., & Archer, J. (1987). Mothers' beliefs about the role of ability and effort in school learning. *Journal of Educational Psychology, 79,* 409-414.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Education Psychology, 80,* 260-267.

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of Educational Research, 64,* 287-309.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44,* 1175-1184.

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp.1-45). New York: Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman & Co.

Bandura, A. (1999). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1-45). New York: Cambridge University Press.

Barron, K.E. & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology, 80,* 706-722.

Biggs, J. (1996). Western misperceptions of the Confucian-Heritage learning culture. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese learner: Cultural psychological and contextual influences* (pp.45-67). Hong Kong: Comparative Education Research Centre.

Butler, R. (1993). Effects of task- and ego-achievement goals on information seeking during task engagement. *Journal of Personality and Social Psychology, 65,* 18-31.

Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: either direction has its ups and downs. *Journal of Personality and Social Psychology, 59,* 1238-1249.

Cervone, D. (1993). The role of self-referenced cognitions in goal setting, motivation, and performance. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction.* (pp. 57-96). Hillsdale, NJ: Lawrence Erlbaum.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934-945.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychology, 41,* 1040-1048.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95,* 256-273.

Dykeman, B. F. (1994). The effects of motivational orientation, self-efficacy, and feedback condition on test anxiety. *Journal of Instructional Psychology, 21,* 114-120.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54,* 5-12.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70,* 461–475.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80,* 501-519.

Fanshawe, J. P., & Burnett, P. C. (1999). Students' and teachers' perceptions of adolescents' problems and coping strategies. *Australian Journal of Guidance & Counselling*, *9*, 93-107.

France-Kaatrude, A-C., & Smith, W. P. (1985). Social comparison, task motivation, and the development of self-evaluative standards in children. *Developmental Psychology, 21,* 1080–1089.

García, J. s.-N., & de Caso, A. M. a. (2004). Effects of a motivational intervention for improving the writing of children with learning disabilities. *Learning Disability Quarterly, 27*, 141-159.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-

regulated strategy development. *Contemporary Educational Psychology, 30,* 207-241.

Hackett, G. (1985). The role of mathematics self-efficacy in the choice of math-related majors of college women and men: A path analysis. *Journal of Counseling Psychology, 32,* 47-56.

Hau, K.T. & Salili, F. (1996). Prediction of academic performance among Chinese students: Effort can compensate for lack of ability. *Organizational Behavior and Human Decision Processes, 65,* 83-94.

Lam, S. F., Yim, P. S., Law, J., S. F., & Cheung, R. W. Y. (2004). The effects of competition on achievement motivation in Chinese classrooms. *British Journal of Educational Psychology, 74,* 281-296.

Levine, J. M. (1983). Social comparison and education. In J. M. Levine & M. C. Wang (Eds.), *Teacher and student perceptions: Implications for learning* (pp. 29-55). New York: Erlbaum.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review, 31,* 313-327.

Marsh, H. W., Kong, C-K., & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effect in Hong Kong schools. *Journal of Personality and Social Psychology, 78,* 337-349.

McAlpine, L. (2004). Designing learning as well as teaching. A research-based model for instruction that emphasizes learner practice. *Learning in Higher Education, 5,* 119-134.

Midgley, C., Anderman, E., & Hicks, L. (1995). Differences between elementary and

　　middle school teachers and students: A goal theory approach. *Journal of Early*

　　Adolescence, 15, 90–113.

Nicholls, J. G. (1979). Development of perception of own attainment and causal

　　attribution for success and failure in reading. *Journal of Educational Psychology,*

　　75, 94-99.

Oettingen, G. (1995). Cross-cultural perspectives on self-efficacy. In A. Bandura (Ed.),

　　Self-efficacy in changing societies (pp.149-176). Cambridge: Cambridge

　　University Press.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational*

　　Research, 66, 543-578.

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics

　　performances: The need for specificity of assessment. *Journal of Counseling*

　　Psychology, 42, 190-198.

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from

　　teacher explanation and repeated listening to stories: Do they overcome the

　　Matthew effect? *Journal of Educational Psychology, 94,* 23-33.

Popham, W. J. (2001). Teaching to the test? *Educational Leadership, 58,* 16-20.

Reyes, L. H. (1984). Affective variables and mathematics education. *The Elementary*

　　School Journal, 84, 558-582.

Richer, S. (1976). Reference-group theory and ability grouping: A convergence of

　　sociological theory and educational research. *Sociology of Education, 49,* 65–71.

Schunk, D. H. (1982). Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology, 74,* 548-556.

Schunk, D. H. (1991). Goal setting and self-evaluation: A social cognitive perspective on self-regulation. *Advances in Motivation and Achievement, 7,* 85-113.

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading and Writing Quarterly, 19,* 159-172.

Schunk, D. H., & Ertmer, P. A. (1999). Self-regulatory processes during computer skill acquisition: goal and self-evaluative influences. *Journal of Educational Psychology, 91,* 251-260.

Schunk, D. H., & Rice, J. M. (1986). Extended attributional feedback: Sequence affects during remedial reading instruction. *Journal of Early Adolescence, 6,* 55-66.

Schunk, D. H., & Rice, J. M. (1987). Enhancing comprehension skill and self-efficacy with strategy value information. *Journal of Reading Behavior, 19,* 285-302.

Schunk, D. H., & Rice, J. M. (1989). Learning goals and children's reading comprehension. *Journal of Reading Behavior, 21,* 279-293.

Schunk, D. H., & Rice, J. M. (1991). Learning goals and progress feedback during reading comprehension instruction. *Journal of Reading Behavior, 23,* 351-364.

Schunk, D. H., & Rice, J. M. (1992). Influence of reading comprehension strategy information on children's achievement outcomes. *Learning Disability Quarterly, 15,* 51-64.

Schunk, D. H., & Rice, J. M. (1993). Strategy fading and progress feedback: Effects on self-efficacy and comprehension among students receiving remedial reading services. *Journal of Special Education, 27,* 257-276.

Schunk, D. H., & Swartz, C. W. (1993a). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18,* 337-354.

Schunk, D. H., & Swartz, C. W. (1993b). Writing strategy instruction with gifted students: Effects of goals and feedback on self-efficacy and skills. *Roeper Review, 15,* 225-230.

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*, 195-208.

Shih, S. S., & Alexander, J. M. (2000). Interacting effects of goal setting and self- or other-referenced feedback on children's development of self-efficacy and cognitive skill within the Taiwanese classroom. *Journal of Educational Psychology, 92,* 536-543.

Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology, 89,* 71–81.

Skinner, E. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology, 71,* 549-570.

Spearitt, D. (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly, 8,* 92-111.

Stevenson, H. W., & Lee, S-y. (1996). The academic achievement of Chinese students. A function of socialization processes and the context of learning? In M. Bond (Ed.), *The handbook of Chinese psychology*, (pp. 124-142). Hong Kong: Oxford University Press.

Thuen, E., & Bru, E. (2004). Coping styles and emotional and behavioral problems among Norwegian Grade 9 students. *Scandinavian Journal of Educational Research, 48,* 493-510.

Urdan, T. (2001). *Statistics in plain English.* London: Lawrence Erlbaum.

Urdan, T. C., & Maehr, M, L. (1995). Beyond a two-goal theory of motivation and achievement: A case for social goals. *Review of Educational Research, 6,* 213-243.

Urdan, T., Midgley, C., & Wood, S. (1995). Special issues in reforming middle level grades. *Journal of Early Adolescence, 15,* 9-37.

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology, 31,* 125-252.

Walker, D., Greenwood, C. R., Hart, B., & Carta, J. (1994). Special children and poverty. *Child Development, 65,* 606-621.

Weiner, B. (1986). An attributional theory of motivation and emotion. New York: Springer.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.) *Self-efficacy in changing societies* (pp.202-231). New York: Cambridge University Press.

Zimmerman, B. J. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22,* 73-101.

Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of Educational Psychology, 89,* 29-36.

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94,* 660-668.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82,* 51-59.

Agent

efficacy expectation

Means   Ends

outcome expectation
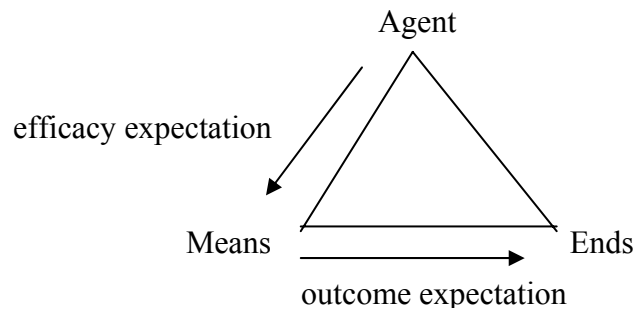
Figure 1. Self-efficacy and the triadic relation among agent, means and ends of control.

Table 1

Means of Self-efficacy by Condition for Study 1

Self-efficacy	Conditions		t value	Cohen's d
	Formative n = 39	Summative n = 40		
Test 1	5.22 [a]	5.37 [b]	-.77	.18
	(.80)	(.87)		
Test 2	4.85 [a]	4.69 [b]	.38	.21
	(.72)	(.80)		
Difference of Tests 1 & 2	.38 [c]	.68 [c]	-.198	.45
	(.54)	(.77)		

Note. The means sharing the same superscript are significantly different from each other. The numbers in parentheses are standard deviations.

Table 2

Means of Self-efficacy by Condition for Study 2

Self-efficacy	Conditions		t value	Cohen's d
	Self-referenced n = 37	Norm-referenced n = 38		
Test 2	4.83 [a]	4.72 [b]	.59	.14
	(.84)	(.78)		
Test 3	5.39[ac]	4.41 [bc]	5.08	1.18
	(.79)	(.89)		
Difference of Tests 2 & 3	.56 [d]	-.32 [d]	6.89	1.60
	(.52)	(.59)		

Note. The means sharing the same superscript are significantly different from each other. The numbers in parentheses are standard deviations.