UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE EFFECT OF FEEDBACK ON SELF-EFFICACY AND MUSICAL APTITUDE SCORES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

PAULA R. HUTTON

Norman, Oklahoma

2013

THE EFFECT OF FEEDBACK ON SELF-EFFICACY AND MUSICAL APTITUDE SCORES

A DISSERTATION APPROVED FOR THE SCHOOL OF MUSIC

BY

_____
Dr. Charlene Dell, Chair

_____
Dr. Michael Raiber, Co-Chair

_____
Dr. Charles Ciorba

_____
Dr. Jennifer Saltzstein

_____
Dr. Teresa DeBacker

Acknowledgements

Researching, organizing, and writing a doctoral dissertation is a mammoth undertaking that could not take place without the personal and academic support of a host of individuals, without whom it is unlikely that the document could be completed. In a phrase recently made popular, it does indeed "take a village" to bring such a comprehensive project to fruition. To all that offered encouragement, said a prayer, or helped in a concrete way, my gratitude is profound.

First of all, I thank Dr. Charlene Dell, who spent countless hours combing through the document, helping with statistics, wording, formatting. She helped in myriad ways to bring the finished product about, and my utmost thanks go to her.

Thanks go to Dr. Michael Raiber, who helped shaped the original research, offering suggestions for sources I had not considered, as well as other insights that shaped the direction of the document. I value his perspective on music education and his dedication to teaching.

I thank Dr. Charles Ciorba for his enthusiastic participation on the committee and his suggestions for the musical aptitude self-efficacy scale.

I thank Dr. Teresa DeBacker for participating as the "non-music" person on the committee. The subject of the dissertation was more in her realm of expertise of educational psychology and her counsel was invaluable.

Special thanks go to Dr. Jennifer Saltzstein for participating on the committee out of kindness. Her milieu is medieval music, not music education research, but she graciously agreed to be part of my committee. Her contribution was her finely honed writing ability, which helped tremendously.

Although Dr. Nancy Barry left OU for a position at another university, I want to acknowledge her help and encouragement. Always tireless in the pursuit of excellence and in helping students achieve their goals, she is the kind of teacher I hope to be. Thank you, Dr. Barry.

Thanks to my music colleagues at Oklahoma Christian University: Dr. Kathy Thompson, Dr. Ken Adams, Dr. John Fletcher, Dr. Heath Jones, and Dr. Kyle Pullen. I appreciate their encouragement and understanding through this journey.

I could not have accomplished any of this without my family, who commiserated with me at times and always cheered me on. My husband Jim learned many new housekeeping skills and always had dinner on the table. My love and gratitude cannot be adequately expressed to him. Our three children and five grandchildren were a never-ending source of encouragement and inspiration. I am so blessed to have this wonderful family.

Finally, ultimate thanks go to God the Father, the Author of everything good. I thank God for music, for His saving grace, and for sustaining me.
To God be all the glory.

TABLE OF CONTENTS

List of Tables

Abstract

The purpose of this study was to investigate the effect of positive and negative feedback on musical aptitude test performance and self-efficacy for musical aptitude test performance. Four research questions were addressed: 1) Would there be statistically significant differences between the scores of participants who have received positive and negative feedback? 2) What would be the interaction effect of feedback and gender on these differences? 3) What would be the interaction effect of feedback and major on these differences? 4) What would be the effect of feedback on participants' self-efficacy for performing on a musical aptitude test?

Participants were 222 university students from two universities in the Central Southwestern United States, one state university and one private university. Edwin Gordon's Advanced Measures of Music Audiation (AMMA) was used as a pretest and posttest. Participants also answered a researcher-designed self-efficacy scale. MANOVA results indicated that feedback did not have a significant effect on the AMMA posttest. However, feedback had a significant effect ($p < .0001$) on self-efficacy levels, with a greater effect on music majors than on non-music majors.

Chapter 1

Introduction

Feedback and Self-Efficacy

The subject of feedback has been important in the fields of education, sports, business, and psychology for decades because it encompasses communication between teacher and student, coach and athlete, employer and employee, scientist and subjects of observation. Feedback has an effect on its recipient, depending on the positive or negative message contained in the feedback and the attitude of the recipient. It has been linked to the construct of self-efficacy, or the belief in one's ability to successfully accomplish a specific task (Bandura, 1977).

Self-efficacy has been erroneously used interchangeably with the constructs of self-esteem and self-concept. However, there are significant differences in their meanings. Self-efficacy, unlike self-esteem or self-concept, concerns one's confidence in being able to accomplish a specific task (Bandura, 1986; Bong, 2006; Bong & Clark, 1999; Bong & Skaalvik, 2003; Gist & Mitchell, 1992). Self-efficacy does not address one's sense of worth or global sense of self. Nor does it ensure success in a given endeavor. An individual may have high levels of self-efficacy toward completing a certain task, but may not feel any motivation to accomplish the task (Bandura, 1986). Thus, self-efficacy is task-specific and motivation-driven.

Self-efficacy is not a static personality trait, but is malleable (Bandura, 1986; Gist & Mitchell, 1992; Vrugt, Langereis, & Hoogstraten, 1997). It is not directly concerned with particular skills, but with the judgments about capable execution of particular skills. The same person may perform a task poorly on one occasion and the same task superbly on another, depending on self-efficacy levels and possible external factors (Bandura, 1986).

Success in learning a new skill or receiving positive feedback can increase an individual's self-efficacy beliefs. Conversely, failure experiences or negative feedback can reduce self-efficacy levels (Bandura, 1977; Gist & Mitchell, 1992). People with high self-efficacy attribute their successes to internal factors such as ability rather than to external causes such as luck. Confidence in one's abilities is not enough to succeed if one lacks the capability to accomplish a task (Bandura, 1977). However, persons with high levels of self-efficacy will persevere when challenged by setbacks, strategizing and increasing effort to reach their goals. Accurate self-appraisal of one's abilities allows an individual to implement the most efficient path to accomplish success (Bandura, 1982a). Successes increase self-efficacy and failures lower it (Bandura, 1977). Therefore, self-efficacy levels can increase or decrease as an individual experiences failures or successes.

In Bandura's model of self-efficacy (1977), four primary sources of information are identified. Each of these informational sources can be viewed

as a form of feedback. An individual will gain knowledge from performance accomplishments, from observing the performance of others (vicarious experience), from verbal persuasion, and from physiological states. Performance accomplishments are the strongest source of information received by the individual (Bandura, 1986). People avoid a performance situation that could end in failure by assessing a task's difficulty and their experience and capability to accomplish the task (Gist & Mitchell, 1992; Silver, Mitchell, & Gist, 1995). A continuous loop exists between feedback, self-efficacy, and performance in which scrutiny of task requirements and attributional analysis help refine the learning process (Gist & Mitchell, 1992). Feedback and self-efficacy interact to influence performance. The current study will investigate the effect of positive and negative feedback, mediated by self-efficacy levels, on performance on a musical aptitude measure.

Vicarious experiences can inform self-efficacy beliefs. Observation of others with similar experience and ability succeeding at a task provides efficacy information to individuals, although vicarious experiences are not as robust as personal performance experiences (Bandura, 1977). Educators provide vicarious experiences for their students by modeling certain success behaviors, both with mastery models and coping models. Coping models demonstrate a person making the same mistakes the students would make in the learning process. Observing another person struggle and then succeed (coping model) is more akin to students' own learning processes than observing a person who

makes no mistakes (mastery model) in the learning process (Bandura, 1997). By observing a model perform, students gather performance cues that include attributions, outcome patterns, and social comparisons (Schunk, 1989).

In addition to learning from vicarious experiences, feedback recipients can learn to attribute failure to external factors rather than to lack of ability (Austin & Vispoel, 1992; Vrugt et al., 1997). When exposed to failure feedback, individuals who subsequently improved their performance did not blame poor performance on internal, unchangeable factors. Instead, they cited lack of preparation and effort as causes. Furthermore, development of strong malleability beliefs can improve performance. Malleability means that ability is not a fixed entity but can be developed through practice, knowledge acquisition, and learning strategies. Individuals who desire successful performance can adjust their attributional style and gain control of external factors that contribute to success.

Bandura's (1977) third source of self-efficacy information is verbal persuasion. To a struggling learner, encouragement from knowledgeable others can help sustain effort and support efficacy levels. The agent of feedback must be credible to the recipient (Gist & Mitchell, 1992; Bandura, 1997). Moreover, the more believable the source of feedback, the more effective is the change in efficacy levels (Bandura, 1986). The deliverer of feedback has the responsibility to give feedback that is realistic about the recipient's capabilities; unrealistic

feedback invites failure and discredits the agent of feedback (Bandura, 1997).

Physiological responses and affective states comprise the fourth source of self-efficacy. Although this component of the informative process pertains more to physical challenges, physiological cues can be important indicators of efficacy levels for non-physical challenges as well. Individuals receive feedback from their own physical responses. Some physical responses are negative; arousal such as stage fright can interfere with successful performance and fulfill its own dire predictions (Bandura, 1997). Thus, individuals gain information about their self-efficacy firsthand through performance, from others in vicarious observation, from verbal persuasion, and from their own physiological responses.

Although by definition, self-efficacy refers to one's confidence in the ability to accomplish a specific task, a general sense of self-efficacy ensues as an individual amasses successful experiences. This sense of achievement generalizes to other situations as the individual gathers skills from experience (Bandura, 1986). General self-efficacy is especially useful for conquering tasks that once were problematic; increases in skills and self-efficacy help individuals replace failure experiences with successes (Bandura, 1977). In academic settings, certain sub-skills can generalize across various domains, so efficacy in one area can transfer to another, similar subject. Bandura (1977) reasons that generalization occurs as individuals synthesize sequences of events, noting the

patterns and actions needed to produce the desired result. It is interesting to note that, whereas task-specific self-efficacy is highly affected by positive and negative feedback, general self-efficacy is not significantly affected by positive and negative feedback (Smith, Kass, Rotunda, & Schneider, 2006). Bandura (1993, p. 118) encapsulates the concept of self-efficacy in this statement:

> The conception of human ability has undergone considerable change in recent years. Ability is not a fixed attribute residing in one's behavioral repertoire. Rather, it is a generative capability in which cognitive, social, motivational, and behavioral skills must be organized and effectively orchestrated to serve numerous purposes. It also involves skill in managing aversive emotional reactions that can impair the quality of thinking and action.

Feedback, Self-Efficacy, and Learning

Teachers provide feedback to students to help them improve their performance. Effective feedback includes remediation, instruction, scaffolding, encouragement, and can increase self-efficacy for learning. Feedback that increases self-efficacy levels may be the most productive type of feedback to provide, because a student's higher self-efficacy helps with self-regulated learning, motivation toward higher goals, and increased self-knowledge (Zimmerman, 2000; Schunk, 1991; Schunk, 2008). However, students interpret feedback through the lens of their prior knowledge and experiences. Feedback

bridges gaps and enriches existing knowledge, corrects misunderstandings, motivates, and increases self-efficacy (Bandura, 1989; Butler & Winne, 1995; Hattie & Timperley, 2007; Shute, 2008; Zimmerman, 2000). Feedback can help diminish learners' cognitive loads by breaking challenging tasks down into manageable steps. Instructive feedback can help teachers scaffold knowledge and skills their students need by providing guidance and efficacy-building encouragement at appropriate times (Margolis & McCabe, 2004; Shute, 2008; Jain, Bruce, Stellern, & Srivastaya, 2007).

Positive feedback interacts with existing self-efficacy levels to affect performance (Karl, O'Leary-Kelly, & Martocchio, 1993). Feedback has a greater effect on individuals with pre-existing high levels of self-efficacy than it does on those with low self-efficacy. Positive feedback provides an incentive to set higher goals among individuals with high levels of self-efficacy. Individuals are more likely to accept feedback that agrees with their pre-existing self-appraisals (Korsgaard, 1996). Bandura (1993) suggests that students with high self-efficacy organize their cognitive tasks more efficiently than students with low self-efficacy. Students with low self-efficacy tend to dwell on their perceived deficiencies, and are slow to recover from setbacks. Feedback to these students must be designed to help them raise their self-efficacy levels and subsequent performance.

Productive feedback increases self-efficacy and improves performance.

Informational feedback aids instruction by helping to bridge gaps in learning, by setting small, attainable goals designed for successful completion, which in turn raises self-efficacy levels. The quality of feedback is important, as well. While positive feedback is preferable, if it lacks specificity, it is not effective. Conversely, feedback that is too specific can stifle creative thinking. Finally, feedback should be structured to raise self-efficacy levels of those who have low self-efficacy, to reverse negative beliefs about their capability, and increase their successes (Hattie & Timperley, 2007).

Positive and Negative Feedback and Self-Efficacy

The quality of feedback, whether negative or positive, has an effect on the self-efficacy of its recipients, and this is especially true of children (Burnett, 1996). Messages children tell themselves about their abilities, or "self-talk," is highly influenced by feedback from significant others in their lives, such as parents, teachers, siblings, and peers. Self-talk reflects self-efficacy levels with statements such as, "I am good at reading," or "I'll never understand math." Teachers in particular influence a child's self-talk regarding academic capabilities. Teachers' positive feedback has a greater effect on girls than it does on boys (Burnett, 1999). Conversely, negative teacher feedback contributes to poor academic performance (Wentzel, 2002). Negative feedback can be interpreted as destructive criticism if it attributes failure to internal causes, such as lack of ability (Baron, 1990). By the time students reach college age, they are

likely to have well established self-efficacy toward particular academic subjects because of the feedback they have received to that point.

Factors can vary in feedback delivery and acceptance. The recipient filters feedback through the lens of existing self-efficacy levels and self-concept. Praise offered for completing familiar, routine tasks can have a negative effect on performance (Waldersee & Luthans, 1994). Individuals who are praised for routine tasks may become self-conscious about their performance, which can interfere with automaticity in performance (Baumeister, Hutton, & Cairns, 1990). In addition to self-consciousness, performers may relax their efforts after receiving praise, resulting in poor performance. However, high and low achievers may receive feedback differently, particularly if it contradicts their performance histories. High achievers may become more focused on achievement after receiving failure feedback, and low achievers may become more focused on achievement after receiving success feedback (Brunot, Huguet, & Monteil, 2000).

In music instruction, there may be a higher proportion of negative feedback to positive feedback (Duke & Henninger, 1998). In a typical instruction scenario, music students receive verbal and non-verbal feedback from their teachers as students work to improve their performance. The teacher creates proximal goals for performance, which invite success or failure feedback as students work to reach these goals. Duke and Henninger's (1998) study

revealed that negative feedback for music students was in the form of directives, which did not affect self-efficacy negatively. The inherent value of improved music performance may have offset any negative feedback effects for students.

Manipulated failure feedback on a standardized musical test may have a detrimental effect on performance among primary school children (O'Neill & Sloboda, 1997). The study suggested that children's self-efficacy levels influenced their performances during testing. The current study will manipulate feedback to university students who take a standardized musical aptitude test. It will be interesting to discover if false positive and negative feedback will have a similar effect on university students.

The relationship of a coach to an athlete is similar to the relationship between a music teacher and a student. Athletes and musicians expect feedback in the form of corrections and encouragement. For this reason, research in music education can glean insights from sports performance research. Many of the performance requirements in music and sports are similarly complex, in that both have physical and psychological components. Athletes may derive increased self-efficacy from the information given to them prior to games (Vargas-Tonsing, 2009; Black & Weiss, 1992). Manipulated failure feedback may influence athletes with high self-efficacy to increase their effort whereas the same feedback may discourage an athlete with low self-efficacy (Weinburg, Gould, Yukelson, & Jackson, 1981). Positive feedback combined with

correction cues for improved performance may increase self-efficacy in athletes (Tzetsis, Votsis, & Kourtessis, 2008). The same could be said of music students receiving similar feedback. The current study will investigate the effect of positive and negative feedback on music students' performance on a musical aptitude test.

Feedback, Self-Efficacy, and Gender Differences

Teacher feedback affects males and females differently, in the way feedback is received and how it affects self-efficacy. Students' processing of teacher feedback indicates gender differences as well (Schmidt, 1995). Positive and negative feedback in the classroom affects males differently than it does females in that males tend to notice teacher disapproval more often than females and females notice teacher approval behaviors more often than males. There is a difference in the acceptance of feedback, depending on the students' gender (Black & Weiss, 1995; Roberts, 1991; Johnson & Helgeson, 2002). Gender differences relating to feedback are present from an early age and continue through adulthood (Beyer & Langenfeld, 2000; Pulford & Colman, 1997).

Women may react to feedback more emotionally than men. In the workplace, women may be more responsive to evaluative feedback than men, regardless of their career level (Johnson & Helgeson, 2002). Women make more behavioral changes after feedback than men, and women are more adversely affected by negative feedback than men. Similarly, female college students are

more sensitive to evaluative feedback than men, in that women regard feedback as more indicative of their abilities than men (Roberts, 1991). College males and females differ in the recall of feedback (Beyer & Langenfeld, 2000). Females tend to overestimate a grade when the feedback is positive and underestimate the grade when the feedback is negative. Women experience more joy after success feedback and more acute disappointment after failure feedback than men, indicating that women may make a more emotional investment in performance outcomes.

There are differences in the way males and females interpret feedback, with women reacting in a more polarized manner than men do, agreeing with feedback, and taking steps to change behavior based on the feedback they receive. Men exhibit more self-confidence in their abilities, and claim internal reasons for success and external reasons for failure, opposite to women's reactions. Since the current study involves giving feedback to male and female college students, a gender effect will be investigated.

Feedback and Test Anxiety

Test anxiety plagues many students, whether the test is cognitively or kinesthetically based. For some, it is an occasional occurrence because of external factors such as lack of preparation, but for others it is a chronic problem. Text anxiety appeared to affect more women than men in a study that investigated the relationship between test anxiety and academic performance

(Chapell, Blanding, Silverstein, Takahashi, Newman, Gubi, & McCann, 2005). Four precursors to test anxiety included purpose of the test, test-takers' confidence about test performance (self-efficacy), familiarity with the test format, and the nature of the material being tested (Reeve, Bonaccio, and Charles, 2008). How test results will be used can also affect test anxiety. When told that test feedback would reveal their IQ compared to other students in the class, university students experienced high levels of test anxiety (Lee, 1999). Similar results were found in another study: test anxiety increased when test-takers anticipated receiving feedback comparing them to other test-takers (Dykeman, 1994). Test-takers experiencing the least test anxiety were those with high self-efficacy who received criterion-related feedback. The relationship of self-efficacy and test anxiety was explored in Pintrich and DeGroot's (1990) study. Self-efficacy was positively related to cognitive engagement and performance and negatively related to test anxiety. Thus, test anxiety can be affected by internal factors, such as gender and self-efficacy levels, and by external factors, such as preparedness, and how test results will be used.

Test anxiety and performance anxiety share many characteristics. Both types of anxiety can interfere with cognitive processes. In a study relating sport and music performance (Tarrant, Leathem, & Flett, 2010), it was found that negative cognitions, negative affect, and irrelevant distractions undermine the performer's concentration and can damage performance. Test anxiety is

significantly associated with lower SAT scores and lower performance on course exams (Cassady & Johnson, 2001). Test anxiety can inhibit performance in anxious students by using a portion of processing capacity needed for concentration (Lee, 1999; Hong, 1999). Students with low test anxiety perform better than those with high test anxiety when informed that the test will be difficult (Weber & Bizer, 2006). Students who tend to be anxious in examination situations can be sabotaged by their own affective responses. Perception of test difficulty, importance of the test results for a course grade, and comparison to other students can exacerbate test anxiety. In addition, test anxiety can interfere with the cognitive processes needed for good test performance.

In the present study, participants will have received manipulated feedback immediately prior to taking the posttest. Feedback will be delivered in such a way as to influence self-efficacy for taking the posttest. Both the pretest and posttest will require considerable cognitive effort from the participants. It is possible that some participants may experience test anxiety as they complete the posttest.

Feedback, Self-Efficacy, and Musical Aptitude

Most researchers agree that musical aptitude is a combination of genetic and environmental influences (Sloboda, 1985; Hodges, 1996; Radocy & Boyle, 2003). Musical aptitude is normally distributed among the population (Gordon,

1989; Hallam & Shaw, 2002; McPherson, 1997). This means that approximately 2% of the population will have high musical aptitude and 14% will have above-average musical aptitude (McPherson, 1997). An individual's musical aptitude stabilizes around age 9 (Gordon, 1989).

Musical aptitude is not one general dimension, but many (Gordon, 1999; Hallam & Shaw, 2002). Components of musical aptitude include constructs that can be tested, such as sense of rhythm, and others that cannot be easily measured, such as emotional sensitivity and commitment to music (Hallam & Shaw, 2002). Thus, it is unlikely that one musical aptitude measure can completely assess an individual's musical aptitude.

Young (1972) compared scores on Gordon's (1965) Musical Aptitude Profile (MAP) between college music majors and non-majors and found that musical aptitude scores for music majors were significantly higher than for non-majors. Gordon later designed the Advanced Measures of Music Audiation test (1989), which is for use with high school and college students. Although the AMMA and the MAP do not measure exactly the same elements, it is important to note that both tests measure musical aptitude and not musical achievement. The AMMA was used to measure musical aptitude of non-music majors in an unpublished study that measured the effect of positive and negative feedback on test scores of non-music majors (Hutton, 2006). It was found that positive and negative feedback had a significant effect on AMMA scores. The current study

will measure the effect of feedback on both music majors and non-majors.

In McCrystal's (1995) study, the AMMA was administered as a pretest in the beginning of the academic year to music majors, and as a posttest at the end of the academic year to the same students. No significant gain in scores occurred from pretest to posttest, indicating that there was no effect of maturation as the students gained knowledge throughout the scholastic year. If no significant improvement occurred in test scores after students had engaged in several months of music study, then it is unlikely that the students in the current study will improve their scores when taking the AMMA twice in quick succession. It also demonstrates that the AMMA tests aptitude rather than achievement.

Need for the Study and Research Questions

The present study will examine the effects of manipulated feedback on musical aptitude test scores and on self-efficacy for performing on a musical aptitude test. As can be seen by previous research, feedback may have profound effects on performance, whether the feedback is genuine or has been manipulated. No published study has investigated the effect of feedback on musical aptitude scores. The current study proposes four research questions:

1. Are there statistically significant differences between the scores of participants who have received positive and negative feedback?

2. What is the interaction effect of feedback and gender on these differences?

3. What is the interaction effect of feedback and major on these differences?

4. What is the effect of feedback on participants' self-efficacy for performing on a musical aptitude test?

It is presumed that the musical aptitude of university students is stabilized and as a result, scores should remain unchanged from one administration of the AMMA to the next. Positive and negative feedback may affect participants' self-efficacy for performing on a musical aptitude test. If feedback has the potential to influence performance on a measure such as a musical aptitude test, then there are ramifications for the importance of feedback in music education.

Chapter 2

Review of Related Literature

Giving and receiving feedback permeates many facets of everyday life, from personal lives to work experiences. There exists a wealth of research on the impact of feedback on the student, the athlete, and the employee. The strength of feedback's effect may be mitigated by established self-efficacy levels and by gender differences. The purpose of the current study is to investigate the impact of positive and negative feedback on musical aptitude scores among music- and non-music majors. Gender and differences between music- and non-music majors on AMMA scores in reaction to feedback will be compared.

Feedback and Self-Efficacy Beliefs

Feedback has been a crucial component of the educational learning process (Duke & Madsen, 1991). Students who strive to achieve often seek guidance from teachers in order to improve performance and learning. A student's self-efficacy, or belief in the ability to accomplish a specific task, can be highly influenced by the person offering feedback. Positive feedback can result in increased self-efficacy and improved performance. Negative feedback, on the other hand, may undermine the student's self-efficacy and the resulting performance may deteriorate. An understanding of how these work separately and together may help teachers better assist their students (Baron, 1988).

According to Bandura, there exist four principal sources of information

from which individuals form a sense of self-efficacy: performance accomplishments, vicarious experience, verbal persuasion, and physiological states. Feedback is overt in verbal persuasion and internal in the other three informational sources. A discrepancy between performance and a personal goal provides feedback information that may affect self-efficacy levels (Bandura & Cervone, 1983).

> The initial approximations of response patterns learned observationally [or by an individual's personal observation] are further refined through self-corrective adjustments based on informative feedback from performance…People process and synthesize feedback information from sequences of events over long intervals about the situational circumstances and the patterns and rates of actions that are necessary to produce given outcomes (Bandura, 1977, p. 192).

Feedback delivery and acceptance plays an integral part in the progression of feedback, self-efficacy beliefs, and performance. Studies by Ayoun (2001), Baron (1988), Amorose and Weiss (1998), and Allen and Howe (1998) indicate that negative or positive feedback can affect self-efficacy levels and resulting performance. Reactions to feedback may differ between men and women. Credibility of the source of feedback may influence the acceptance of feedback; individuals are more likely to accept feedback from a perceived

expert. The effect of feedback is persuasive enough to affect test performance (Bouffard-Bouchard, 1990).

The Effect of Positive and Negative Feedback on Self-Efficacy

In a study of negative and positive feedback effect in a French language class, Ayoun (2001) suggested that implicit negative feedback in written form was more effective than explicitly negative verbal correction. The study compared three methods of language correction: implicit negative feedback in written form (called recasts), preemptive positive evidence (models), and traditional classroom grammar instruction, which involves both positive evidence and negative feedback. Participants were 145 French language students. A pretest was administered in which participants provided corrections to 24 sentences that featured various syntactic errors. Next, participants completed a short written assignment in which they were asked to include certain grammatical elements. The total scores from the pretest and the written assignment determined the overall raw score that was used to place participants in one of three levels of grammar competency: low, middle, or high. Participants in each level were randomly assigned to one of three treatment conditions: traditional grammar instruction, recasting, or modeling. Participants in the traditional grammar group received positive, explicit evidence in the form of traditional grammar being modeled. They were encouraged to read the answer key upon finishing each grammar session. This constituted explicit negative

feedback for incorrect answers. Recast and model group participants read a different story with illustrations each week. Participants in the recast group were asked to form a sentence with elements of the story and to conjugate verbs used in the story. Participants in the modeling condition were shown a sentence corresponding to the illustration for three seconds as pre-emptive positive evidence and then answered a related question. Last, participants from all groups composed another short written piece as a posttest so that gains from pretest to posttest could be measured. Results of a one-way ANOVA revealed no differences between groups on the pretest, indicating that participants were initially at the same level of competency. A second ANOVA indicated that all three groups improved significantly from pretest to posttest ($p < .001$). The recast group showed the most improvement (+38.74%) of the three groups from pretest to posttest. The modeling group moderately improved (+32.62%), and the grammar group improved the least (+27.91%). Learning occurred in all three conditions, but participants who received implicit negative feedback outperformed participants who received explicit positive evidence and negative feedback found in traditional instruction.

Music instruction is similar to language instruction in that it involves giving corrective, or negative, feedback to improve performance. In a study designed to compare the effects of negative feedback and specific directives in music instruction (Duke & Henninger, 1998), 25 college undergraduates and 25 fifth- and sixth-grade students received instruction on playing a soprano

recorder. The same teacher taught all participants individually. In approximately half of the lessons for each age group, the teacher used corrective (negative) feedback. The other group heard specific, non-corrective directives only. Descriptive statistics revealed no significant affective or performance differences for the two treatment groups ($p > .30$). All participants reported having a positive experience. The time required to reach the performance goal was not different between the group receiving negative feedback and the group receiving directives. The results also indicated that the effects of negative feedback were mitigated by the goal of performance achievement, rendering additional positive reinforcement unnecessary.

Feedback can have an effect opposite of the one intended. Baumeister, Hutton, and Cairns (1990) included four experiments in their study, testing male undergraduates playing an Atari video game. The number of participants varied in each of the four experiments, from 28 to 62. ANOVA statistics from each experiment revealed a significant negative effect of praise on performance ($p < .02$ to $p < .001$), especially for praise given immediately before performance. Results indicated that praise appeared to increase effort, but it may have impaired performance by pressuring the performer to do well and by increasing self-consciousness. Self-consciousness can disrupt automatic processes necessary for a skilled performance. Praise may convey the expectation for continued skilled performance; it also may cause a relaxation of

effort from the implication of the goal being met. Thus, praise intended to improve performance might have the opposite effect.

Negative feedback can take the form of destructive criticism, or feedback that violates the goals of being specific and considerate, instead attributing poor performance to causes within the individual rather than to external, changeable causes (Baron, 1988). In a three-part study investigating the effects of destructive criticism in the workplace, employees rated destructive criticism as an important cause of conflict, especially when a peer or subordinate delivered it. Such criticism also had an effect on task performance and goal setting by reducing efficiency and accuracy. The feedback was manipulated to be constructive for one group and destructive for a second group. A third group received no feedback. The three groups completed clerical and proofreading tasks. Data measured by analysis of variance revealed a significant main effect ($p < .0001$) for mode of criticism. Not surprisingly, only the group receiving destructive criticism reported reduced self-efficacy, especially in proofreading accuracy. Results suggest that destructive criticism lacks consideration for the recipient, and that attributing poor performance to internal causes will not bring about improvement. In fact, such criticism may have the opposite effect.

Reluctance to accept negative feedback was the focus of a study investigating the moderating role of self-efficacy on feedback acceptance (Nease, Mudgett, & Quiñones, 1999). Eighty undergraduate psychology

students played a computerized Naval Air Defense simulation. Before each trial, participants completed a self-efficacy questionnaire. After each trial, participants received preselected, manipulative feedback followed by another questionnaire asking about acceptance of the feedback. A 2x2x3 mixed-factor design was used with self-efficacy (low vs. high), feedback sign (positive vs. negative), as between-subjects factors, and time (three trials) as a within-subjects factor. Results indicated that individuals with high self-efficacy were significantly less likely to accept negative feedback ($p < .001$) than positive feedback, even though they received three times as much negative feedback as positive feedback. Participants with lower self-efficacy were more likely to accept negative feedback than those with higher self-efficacy. Results of this study indicated that acceptance of positive or negative feedback was mediated by pre-existing self-efficacy beliefs.

Positive feedback can increase self-efficacy levels, which may influence taking risks in decision-making (Krueger & Dickson, 1994). Participants were 153 business majors who underwent a first round of decision-making that involved hypothetical business risks. Next, participants reported their levels of self-efficacy in regard to the perceived opportunity for success and threat of failure in making the first round of decisions. Participants were randomly assigned to receive false feedback on their decisions for each of two tasks. After the second round of decision-making, participants again reported self-efficacy levels and perceptions of threat of failure or opportunity for success.

Correlational analysis and analysis of variance indicated a significant effect of feedback on self-efficacy. Positive feedback increased self-efficacy while negative feedback decreased self-efficacy ($p < .0001$). Perceived opportunity and subsequent risk-taking correlated positively to increased self-efficacy. Results from this study indicated that participants who were led to believe they were competent decision-makers noticed more opportunities in a questionable choice and took more risks. Participants who were led to believe that they were not good decision-makers saw situations as more threatening and took fewer risks. Increased self-efficacy in decision-making did not generalize to other tasks.

In summary, positive and negative feedback can take many forms. Feedback can be verbal or written, and it can be in the form of praise or overt criticism. The most common effect of positive feedback is improved performance, and the effect of negative feedback most often is increasing failure, although there are exceptions. Praise may raise an individual's level of self-consciousness and increase pressure to perform well, thereby thwarting successful performance by diverting attention from the performance to the performer. Feedback that contradicts an individual's established self-efficacy is likely to be rejected; conversely, feedback that agrees with established self-efficacy levels is more likely to be accepted. Positive feedback has been found to increase risk-taking in decisions and negative feedback has been found to inhibit decision-making. The current study will employ both positive and

negative feedback to participants. Feedback will inform half of the participants that they had high scores on a musical aptitude test, while the other half of the participants will be informed that their scores were very low. The research above suggests that some participants in the present study who have high self-efficacy for music may reject the negative feedback message altogether, or may succumb to its negative implication that they have low musical ability. Conversely, participants who have low self-efficacy for music may be encouraged by the positive feedback message and give more effort to perform well on the posttest.

Feedback, Self-Efficacy, and Gender Differences

Research suggests that there may be differences in the manner in which males and females accept performance feedback. Steinmayr and Spinath, (2009) measured boys' and girls' intelligence and parental stereotypical perceptions of their children's intelligence to discover reasons behind boys' stronger confidence in their own intelligence. Participants were 496 high school juniors and seniors from Germany who estimated their own intelligence. Parents also rated their children's intelligence. Participants were administered the Intelligence Structure Test (Amthauer, 2001, as cited by Steinmayr & Spinath, 2009). The test evaluated domain-specific intelligence in these areas: verbal, numeric, figural, and reasoning. Parental ratings of their children's intelligence were evaluated using a measure designed by Steinmayr and Amelang (2006, as cited by

Steinmayr & Spinath, 2009). They were asked to rate their children's verbal, numeric, figural, and reasoning abilities. Participants were asked to rate their own intelligence using the same scale. Results of a MANOVA, with child's gender as the factor, and measured intelligence, intelligence self-estimates, and parental judgments as dependent variables revealed a significant main effect for child's gender ($p < .001$). Multiple ANOVA results indicated that boys rated their numerical, figural, and reasoning intelligence higher than girls ($p < .001$). Parents also rated these intelligences higher for boys than for girls ($p < .001$). Boys outperformed girls not only on the numeric ($p < .001$), figural ($p < .05$), and reasoning intelligence ($p < .001$), but also verbal abilities ($p < .01$). Regression analyses were performed to understand whether the gender effect on intelligence self-estimates could predict differences in measured intelligence. A main finding of this study was that boys' stronger confidence in their intelligence and their parents' higher ratings only partly attributed to boys' superior performance on intelligence tests. It was surmised that boys' and their parents' beliefs in boys' higher intelligence influenced the results whereas girls and their parents believed that girls would perform less successfully which resulted in lower performance from girls.

Anchoring bias is the persistence of self-efficacy beliefs when faced with results to the contrary. The effect of anchoring biases and the perseverance of self-efficacy beliefs when receiving performance feedback revealed a gender difference in a study by Cervone and Palmer (1990). Participants were ninety-

six university psychology students randomly assigned to one of four anchor bias conditions: Low Anchor, Intermediate Anchor, or High Anchor, or No Anchor (control) group. Participants were given 30 "cyclical graphs" to solve by tracing over lines of designs without lifting the pencil or retracing any lines. Success was manipulated by controlling the number of graphs that could be completed. The final 13 of the 30 graphs were unsolvable. Participants judged their exact level of self-efficacy on questionnaires before each subset of 10 graphs by indicating the number of graphs they believed they could complete successfully. After completing all 30 graphs, participants completed a posttest questionnaire reassessing their self-efficacy levels. Results indicated that men initially judged they could solve a higher percentage of items than women ($p < .01$). Despite less success as the trials progressed, men still had higher judgments of their capabilities than women. An ANOVA indicated that the anchoring manipulation significantly affected initial self-efficacy judgments ($p < .0001$). High, intermediate, and low anchor values, though induced by manipulation, generated high, intermediate, and low initial levels of self-efficacy. Men had higher perceptions of their abilities than women in the last 20 task items ($p < .01$). These results suggest that anchoring biases may be persistent in spite of performance to the contrary and that males tend to rate themselves higher than women do, even after failure experiences for both genders.

Gender differences were discovered in a study of causal attributions after taking an exam (Beyer, 1998). Participants were two hundred forty-six

undergraduates (156 females, 90 males) enrolled in a general psychology course. They were told to imagine receiving either an A or an F in a course examination and to state the reasons for either success or failure; the exam was important to their graduation. Course names and numbers of actual classes taught at the university were used to create realism. Participants ranked a list of possible reasons for making an A or an F on an exam. They also rated how likely they were to make the same grade on another exam, and whether they were responsible for the grade or whether there was an external reason, such as bad luck or lack of study. ANOVA results revealed a significant main effect for gender ($p < .03$), indicating that males ranked and rated ability as a more important cause of success than did females. Females were more likely than males to admit feeling like a failure. Males tended to protect their self-confidence in failure situations by blaming a poor performance on an unstable cause that could be changed, whereas females attributed failure to an unchangeable cause such as ability. After success, females reported feeling prouder, happier, and more confident than did males. The researchers point out, however, that the genders are more alike than they are different in causal attributions, with both genders agreeing that preparation and study are more important to success than luck.

Johnson and Helgeson (2002) investigated gender differences among 90 employees in responsiveness to evaluative feedback. Bank employees' annual evaluations took place in a face-to-face manner. Evaluations were linked to

salary increases. Researchers hypothesized that the effect of feedback would be greater for women than for men, both to their self-esteem and to post-feedback behavior. Other components of gender differences toward evaluation were investigated: whether women had lower expectancies of the evaluation outcome, whether women found the evaluative feedback to be more informative than men, whether women took the feedback more seriously than men, and whether men and women regarded the credibility of the supervisor differently. Participants answered one questionnaire before the evaluation and a second questionnaire shortly after the evaluation meeting. On the first questionnaire participants indicated the score they expected to receive on their yearly evaluation, based on a 7-point scale used by the employer. On the second questionnaire participants rated their response to their evaluation using a 5-point scale as to how they perceived the score, from "very negative" to "very positive." In addition, participants indicated how much they agreed with the positive and negative comments made during their evaluation. On both the first and second questionnaires, participants indicated their satisfaction with the evaluation process, whether the evaluation was personally helpful to them, whether they regarded the evaluation process as fair, and whether they took the evaluation seriously.

After completing the two questionnaires, participants were administered the Rosenburg Self-Esteem Scale (1965, as cited in Johnson & Helgeson, 2002), to measure "situation-specific self-esteem" (p. 246). Results from multiple

regression analyses indicated that gender and feedback were a marginally significant prediction of self-esteem ($p < .001$). Positive and negative feedback had very little effect on men's self-esteem. Women's self-esteem, however, increased slightly after positive feedback but declined dramatically after negative feedback. Simple slope analyses indicated that the slope for feedback effect was significant for women (*beta* = .25*; p* < .005) and not significant for men (*beta* = .08). Regression analysis indicated that women were more likely than men to make changes in their work behavior after feedback ($p < .05$). It should be noted that although the component being measured in this study was self-esteem and not self-efficacy, the finding of feedback effect for gender was significant to the present study. It could be argued that the construct "situation-specific self-esteem" mentioned in the study was actually self-efficacy, as self-esteem is a more general self-construct whereas self-efficacy by its definition is situation-specific.

Reasons individuals cite for failure or success is called "explanatory style." Seligman, Nolen-Hoeksema, Thornton, and Thornton (1990) conducted two studies investigating explanatory style after manipulated failure feedback was given to varsity men's and women's swim teams. In the first study, the Attributional Style Questionnaire (Peterson, Semmel, von Baeyer, Abramson, Metalsky, and Seligman, 1982) was administered to swimmers: 21 men and 26 women at the University of California at Berkeley to discover explanatory style for good and bad events. Regression analysis results suggested notable

differences in explanatory style between genders: Men had greater optimistic explanatory style than did women ($p < .001$). In the second study the scorekeeper imposed failure on the swimmers with a false slow score and tested their performance following defeat. The slower scores were low enough to induce defeat in the swimmers but not so low to raise their suspicions. Swimmers with an optimistic explanatory style for negative events exhibited only marginal decline in the mean ratio of their scores for swim time (M = -.094), but performance worsened more dramatically among swimmers with a pessimistic explanatory style (M = -.833). The researchers stressed that this was a dramatic difference in swim times. The two studies suggest that an optimistic or pessimistic explanatory style may influence performance after defeat, and that men tend to exhibit a more optimistic explanatory style than do women. Thus, men accept negative feedback differently than women.

Men's optimistic explanatory style may manifest itself as overconfidence to perform well on a general knowledge test. Pulford and Colman's (1997) study investigated the effects of everyday types of feedback on confidence levels between genders, whether confidence levels will adjust after feedback to match accuracy level. Participants were 48 male and 102 female undergraduate students randomly assigned to six treatment groups. Twenty general knowledge questions were drawn from the game Trivial Pursuit with three levels of difficulty: hard, medium, and easy. Questions were read and participants wrote down their answers and their level of confidence that their answers were correct.

Some groups received feedback after each question was answered and some groups received no feedback until all they had answered all the questions. Descriptive statistics suggested that no significant effects were found between the feedback groups and the no-feedback groups for accuracy or confidence. However, males were significantly more overconfident than females (males M = -0.74 vs. females M = -6.23, $p < .01$). Males also found the questions easier than did females at all levels of difficulty, and achieved higher accuracy scores than females did (males M = 59.06% vs. females M = 49.43%, $p < .001$). Both male and female groups responded with overconfidence but females were significantly less confident and accurate than males for all levels of item difficulty. Because females believed they were faring worse than the males, based on the negative feedback they received, their confidence levels dropped accordingly, and their subsequent performance suffered.

Gender differences in mediating music teacher feedback were investigated among choral students (Schmidt, 1995). Participants were one hundred twenty secondary school choral students attending a choral summer camp. They listed their reasons for success or failure in music; their answers were coded using Weiner's (1974) Attribution Theory model. Next, participants rated 29 audio-taped excerpts of vocal instruction. The short excerpts featured a teacher giving approving or disapproving comments to a singer. Categories to rate each excerpt included such items as "good/bad," "sincere/insincere," "effective/ineffective," etc., and ratings were coded. ANOVA results indicated a

significant effect ($p < .001$) between genders. Females' ratings of teacher approval were higher than males' ratings, across the grade levels. Thus, student gender appears to mediate teacher feedback in music instruction, with females giving higher approval ratings of teacher feedback than males.

Just as there are gender differences in causal attributions, there also may be differences in the recall of performance feedback. In a study by Beyer and Langenfeld (2000), participants (88 female and 68 male university students) were assigned to read evaluative feedback given for an English paper and a computer program. They were asked to guess the letter grade that had been assigned to the paper and the program, and to rate the feedback. A three-minute surprise test was administered in which participants were asked to recall the positive and negative performance feedback. ANOVAs revealed no gender effect for the computer program feedback, but a significant ($p < .006$) gender effect was found for the English paper: Women had more polarized reactions than did men. When the feedback was positive, women estimated higher grades than did men; when feedback was negative, women estimated lower grades than men did, indicating that women may be more affected by evaluative feedback than are men.

Research suggests that there are gender differences in reaction to feedback. Women appear to be more vulnerable to its effects, and are more concerned with making changes after receiving feedback. Men, on the other

hand, feel more confident in their intelligence and abilities, have a more optimistic explanatory style than women, and tend not to accept feedback to the extent that women do. The present study will focus on gender differences in accepting feedback.

Self-Efficacy and Manipulated Feedback

Manipulated feedback induced high- or low self-efficacy in participants in a French class (Bouffard-Bouchard, 1990). Results suggested that self-efficacy operated partially independently of requisite skills for speaking and reading French. Participants were 64 Canadian students enrolled in a university French class. High and low self-efficacy was induced in two randomly assigned groups of participants. Each target word was concocted and had no real meaning other than the context provided by the sentence. Each student indicated a level of confidence that choices were correct on four problems of varying difficulty. Verbal feedback to individuals in the induced high self-efficacy condition was positive and emphasized success in comparison to others in the study. Verbal feedback to the induced low self-efficacy condition was negative and reported failure in comparison to others in the study. To measure self-efficacy levels, students were presented with four more problems and asked whether they believed they could solve them. Students also indicated their level of confidence in providing correct answers. Students from the high self-efficacy condition completed significantly more problems than did the group with low

self-efficacy condition. The high self-efficacy group worked the problems more carefully and thoroughly than did the low self-efficacy group, resulting in better scores. ANOVA results indicated that those with high self-efficacy beliefs were significantly more accurate in evaluating their responses ($p < .0005$) than were students with low levels of self-efficacy because they were more thorough and careful, and were more confident of their choices.

The effect of manipulated failure feedback on self-efficacy and subsequent performance was the focus of two studies involving students committed to professional goals (Brunstein & Gollwitzer, 1996). Both experiments challenged students' identity as professionals in their chosen fields. Participants in the first study were 96 medical students. It was explained to the participants that being able to pay close attention is characteristic of doctors. Participants were assigned a mental concentration task that demanded high levels of effort and close attention. Descriptive data were gathered from this task. Participants also self-reported their levels of motivation and preoccupation with failure. In the next phase of the experiment half of the participants were assigned to a no-feedback control treatment and the other half received manipulated failure feedback. In the second task the students were introduced to a social problem a doctor might encounter. Then participants were instructed to choose a solution to the problem from four possible solutions provided. Next, participants were assigned a concentration task, with the suggestion that successful doctors must be able to concentrate. The concentration task involved

a computer-generated letter cancellation game that had to be worked as quickly as possible. Finally, participants answered questionnaires to measure the effect of the failure manipulation. On the first task, involving mental concentration, a MANOVA revealed a significant effect of failure feedback ($p < .01$). Results of the second task, in which participants chose a solution to a social problem, a significant effect was found from the failure feedback of the first task ($p < .001$) for the manipulated-feedback group, but no significant effect for the non-feedback group ($p < .10$). Participants reported being beset with memories of failure from the first task that interfered with concentration on the second task. Thus, when the participants' professional identity was challenged, failure feedback may have carried more importance because it reflected potential professional shortcomings. The current study will involve music majors who will receive manipulated feedback that may challenge their identity as accomplished musicians.

The second phase of the above-mentioned study (Brunstein & Gollwitzer, 1996) again involved manipulated failure feedback. Similar to the previous experiment involving medical students, the second experiment involved ninety university students who were committed to a future in computer science. This experiment, however, differed from the first experiment in that it concerned restoring pre-existing self-efficacy to feedback recipients. The experiment was comprised of three phases: treatment task involving manipulated feedback, a completion-inducing intervention for the participants

who had been exposed to manipulated failure in the first experiment, and a final test task. A 3x2, Pretreatment x Test factorial design was used. The pretreatment factor consisted of three conditions: no feedback, failure only, and failure followed by feedback designed to restore self-efficacy for computer science. For the treatment task, participants took a computerized concept formation test that contained eight questions in which participants were to predict the experimenter's choice from eight options (Brunstein & Olbrich, 1985). The test task consisted of the d2 Mental Concentration Test (Brickenkamp, 1981, as cited by Brunstein & Gollwitzer, 1996) that tested visual alertness. Participants were informed that the test measured skills important to computer science. While taking the test, one-third of participants received no feedback, and the other two-thirds received failure feedback on their computer screens as the test progressed. A pattern of initial successes was preprogrammed, followed by a period of continual failure. Feedback for each question was immediate and appeared on the participants' computer screen. After completing the test, all participants completed a personality questionnaire designed to reveal characteristics of successful computer scientists. Half of the failure-feedback group was assigned manipulated personality feedback designed to restore computer-science self-efficacy after failure. They were led to believe they shared characteristics with highly qualified computer scientists. Results of an ANOVA revealed higher scores for the group receiving false positive feedback than for the no-feedback group ($p < .001$). This study indicates that

effects of failure on an identity-relevant test may be reversed when a new appraisal of identity-relevant skills matches participants' self-descriptions.

Not only does the quality of feedback have an effect on the recipient of feedback, but the source of the feedback also plays a part in the effect of the feedback on the recipient (Bannister, 1986). It was hypothesized that feedback is more likely to be accepted if the feedback recipient believes the source is credible. Bannister experimentally manipulated characteristics of the feedback message and the source of the feedback. Undergraduate business students (N=149) participated in the study. A 2x3x2 factorial design was used that measured three components: favorableness of initial performance outcome, attributional feedback, and credibility of feedback source. Participants were randomly assigned to one of 12 treatment groups. They performed a task involving truck routing and received performance feedback and bogus percentile ratings. Participants answered a questionnaire assessing performance attributions. Evaluators' credibility also was manipulated: one evaluator was a vice president with years of supervisory experience, whereas the other evaluator was described as a 22-year-old college sophomore and part-time intern who allegedly had worked on the truck routing problem but not successfully. Results from an ANOVA revealed a significant main effect between the positive and negative outcome conditions ($p < .0001$), indicating that participants who received initial positive feedback assumed more responsibility than those who received negative performance feedback. The credibility of the source of the

feedback also had a significant effect. Participants receiving feedback from the credible source judged the feedback as more accurate ("credible" = 4.95 vs. "not credible" = .4.34) using a 7-point Likert-type scale. Participants regarded the source as more perceptive and they expressed greater satisfaction with the feedback. Thus, both the message and the messenger are important when feedback is delivered.

General self-efficacy is an individual's confidence in being able to solve a broad range of problems, as opposed to task-specific self-efficacy, which is an individual's confidence in the ability to solve a specific problem. To explore the possible difference between feedback's effect on general and task-specific self-efficacy, anagrams were assigned to 60 university students (Smith, Kass, Rotunda, and Schneider, 2006). Participants were divided into control and experimental groups. Both groups solved a series of anagrams of average difficulty, after which feedback was given. The experimental group was given an additional set of anagrams, five of which were difficult and five that were unsolvable. General and task-specific self-efficacy was measured before and after the experiment using the New General Self-Efficacy Scale ([NGSE]; (Chen, Gully, & Eden, 2001) and the Task-Specific Self-Efficacy Scale, comprised of items drawn from the NGSE by the researchers. Results from a mixed-design ANOVA revealed a significant effect ($p < .01$) of failure feedback on task-specific self-efficacy but no effect for general self-efficacy, suggesting

that general self-efficacy and task-specific self-efficacy are not affected in the same way by failure feedback.

As with other academic tasks, performance on a musical test requires cognitive processing (O'Neill & Sloboda, 1997). In a study that focused on the effects of test failure on children's musical test performance, 51 children, aged 6-10, were administered a standardized Melodic Direction Test from the Measurements of Music Listening Skills (Simons, 1976). Next, the children took a similar, researcher-devised test that also measured recognition of melodic direction, on which the children were rated as successful. A third test resulted in experimenter-induced failure by requiring the students to indicate one direction of each melody when some of the melodies changed directions. After the second and third tests, the children were asked to report their self-perceptions of musical competence and to predict their performance on future tests. After experiencing failure, the performance deteriorated in more than half the children. Descriptive statistics revealed a drop in confidence levels for all age groups after failure feedback, and significant decreases for seven- and eight-year-olds ($p < .001$), though the reason for the effect on this age group was not explained. The children who had high musical self-efficacy following failure tended to maintain or increase their performance on subsequent tests, but children with low levels of musical self-efficacy showed poorer performance. The study demonstrates that, though music cognition is dependent on cognitive skills, self-efficacy beliefs also play an important role in musical performance.

Although failure feedback following performance can have an effect on subsequent performance, pretest feedback given immediately before retesting can have an effect on subsequent test performance as well (Bridgeman, 1974). False scores from a scholastic aptitude test were given to 233 seventh-graders immediately before they took a test comprised of items from the non-verbal battery of the Lorge-Thorndike Intelligence Test. Bridgeman hypothesized that students given success feedback on the pretest would score higher on the posttest than those who received failure feedback, and that there would be a gender effect from the pretest feedback. Although a gender effect was not found, results from regression analysis indicated that students, regardless of gender, who received high pretest score manipulation scored significantly higher than students who received low scores ($p < .01$). Results indicated that students' perception of prior success or failure influenced subsequent test performance. Interestingly, in the thirty-plus years since Bridgeman's study was published, no replications of this study have been conducted. However, it shares two important details in common with the current study: use of a standardized aptitude test, and feedback manipulation of pretest scores.

Self-efficacy can be manipulated by feedback, as the studies above indicate. Failure feedback can have an especially detrimental effect on individuals when it implies lack of aptitude for a chosen career. Manipulated positive feedback can encourage greater effort to succeed in a task, whereas manipulated negative feedback can discourage effort. Self-handicapping may

occur if individuals are hesitant to know the true extent of their abilities. In addition, warning of a test's difficulty immediately before administering the test may intimidate test-takers. The current study will manipulate feedback to participants in an attempt to affect participants' self-efficacy levels regarding their posttest performance.

Feedback, Self-Efficacy, and Musical Aptitude

It is generally accepted that all individuals possess musical capacity to some degree, but opportunities to develop musical capacity vary greatly among the population. In Hallam and Shaw's (2002) qualitative study, participants completed the statement "Musical ability is…" (p. 104). Responses were coded into categories of musical ability that included Musical Ear; Rhythmic Ability; Listening and Understanding; Response to Music; Being Able to Play or Sing; Reading Music; Declarative Knowledge about Music; Technical Skills; and Emotional Sensitivity. Participants responded to statements in each category by using a 5-point rating scale. Children and teachers from a city school, an extra-curricular music school, and a higher education institution took part in the study. The youngest participants were 14 years old. Some participants were musicians with varying degrees of musical involvement and others had no involvement with music. An ANOVA was performed to explore differences between groups of participants and their responses. No significant differences were found between respondents for these categories: musical ear, being able to play or sing,

rhythmical ability, listening and understanding, response to music, or being able to read music. Significant differences were found for all other categories, some of which were: emotional sensitivity ($p < .0001$), personal expression ($p < .008$), composition and improvisation ($p < .003$), motivation ($p < .004$), and generative activities ($p < .0001$). All the children exhibited high response rates for appreciation of music, evaluation, and knowledge. Musicians indicated that musical ability was related to communication, ensemble skills, emotional sensitivity, and organization of sound. Results indicated that there was more agreement between children and two teacher groups over particular musical aspects than there was disagreement. However, differences in defining the nature of musical aptitude reflected the degree of active musical involvement of participants. Children with limited experience in music gave highest ratings to appreciation and knowledge of music, and to improvisation and musical creativity. Children with musical experience responded most to music appreciation and creativity. Musicians emphasized skills beyond singing or playing an instrument, e.g., ensemble skills, music as communication, musical sensitivity. Non-music educators preferred to define musical ability as a way to express emotion through musical means. These viewpoints demonstrate a basic difference between musicians' and non-musicians' definition of musical ability. The current study will investigate musical aptitude in university music majors and non-majors. There may be differences between the groups in their definition of musical ability, as there were in the Hallam and Shaw study. These

differences between groups may influence their performance in the current study. Non-musicians may not be concerned about performing well on a musical aptitude test if musical aptitude is not important to them whereas participants for whom music is important may exert more effort on the same test.

Young (1972) conducted a study in which Gordon's (1965) Musical Aptitude Profile (MAP) served two purposes: first, to investigate any differences between MAP scores of freshmen music majors from Midwestern universities and scores of music majors from southern universities; and second, to establish MAP norms for use with non-music major college freshmen. Participants were 205 elementary education majors enrolled in a music fundamentals course. Most participants had not received any formal music training. The MAP battery was administered early in the semester. Raw scores were transformed into standard scores; scores of the participants were compared with standard scores for high school students and freshman music majors. Standard score results indicated that, without exception, means for non-music major university students were considerably below those for university music majors. With 100 representing a perfect score, results indicated that the standard score mean for university music majors was 64.0 compared to the mean score for non-music majors: 57.5, and the mean score for high school seniors in general, which was 53.9. There were no substantial differences in standard scores between freshman music majors from Midwestern and Southern universities. The researcher stressed that

musical aptitude, as measured by the MAP instrument, may be comparable for university students in any area of the country. The current study will use Gordon's (1989) Advanced Measures of Music Audiation (AMMA). While both the MAP and the AMMA test musical aptitude and either test can be used with university students, the AMMA will be used in the current study. The MAP is designed for use with students in fourth grade through college age, but the AMMA is for testing high school seniors and college students. Furthermore, the AMMA takes less time to administer than the MAP (AMMA = 20 minutes; MAP = 150 minutes) and provides valid results.

Non-music majors were participants (N = 101) in an unpublished study investigating the effects of feedback on musical aptitude test scores (Hutton, 2006). Participants were enrolled in a music appreciation class; some had formal music training and others had none. Gordon's AMMA (1989), was used because of its high validity and reliability. The AMMA was administered as a pretest and as a posttest to gain raw scores from participants. To gauge the effect of feedback on self-efficacy and subsequent test performance, manipulation occurred immediately prior to administration of the posttest. Designations of "A" for high pretest scores and "B" for low pretest scores were randomly assigned to participants and written on their posttest answer sheets. The AMMA was administered again, this time as a posttest. ANOVA analyses revealed a significant main effect from the manipulation of pretest scores to posttest performance (p = .02). Scores between Group A and Group B were

significantly different for the posttest (p =.01). Scores of participants who had prior music experience were significantly higher than scores of participants with no prior music experience (p = .02). There also was a significant effect for differences between genders. With 100 representing a perfect score, mean scores for males were higher than those for females on the pretest (Male M = 51.43 vs. Female M = 46.63), but females scored higher than males on the posttest (Female M = 54.15 vs. Male M = 53.15), indicating that females may have been significantly more receptive to positive feedback from the pretest. Results suggest that performance on a musical aptitude test may be significantly affected by positive or negative feedback.

Summary

From Bandura's (1977) original work with self-efficacy, research has investigated the effect of positive and negative feedback on self-efficacy and performance. Existing self-efficacy moderates acceptance of feedback and influence resulting performance. Individuals with high self-efficacy tend to perform at higher levels than those with low self-efficacy whereas those with low self-efficacy tend to self-handicap.

There are gender differences in reaction to feedback. Males tend to overestimate their abilities, cite external reasons for failure, and discount feedback that disagrees with their own perceptions. Conversely, females tend to react more strongly to feedback, citing internal causes for failure. Manipulated

positive and negative feedback can have significant effects on behavior. False positive feedback may result in increased performance and false negative feedback may influence individuals to doubt their abilities to perform well in their chosen careers.

Not all positive feedback has a positive effect. Receiving praise may influence recipients to become self-conscious about their performance or to relax their efforts. Individuals studying music are likely to receive feedback in the form of praise, but students will generally receive more negative feedback in music than feedback in other areas. Negative feedback in the form of corrective statements is commonly accepted in music study and it is thought that the reward of improved music performance outweighs the effects of teachers' negative feedback. There are similarities between music and sports, as both are performance-based disciplines. Thus, much of the research into feedback, self-efficacy, and performance has benefit for both areas.

The current study will address a key aspect for which there is existing research: how feedback affects performance. However, very little research has been conducted on the effect of feedback on music test scores and even less research exists on the effect of feedback on scores of a musical aptitude test. Factors of this study include attributions of success or failure, the effect of positive and negative feedback, and cognitive processes in music as participants perform on a musical aptitude evaluation

Chapter 3

Research Design and Methodology

Feedback's effects have been of interest to researchers and educators for the past several decades because of the very important role feedback plays in the learning process. Self-efficacy, or one's level of confidence to perform a specific task, has been the focus of research since Bandura (1977) introduced the construct. Feedback often interacts with an individual's self-efficacy to effect change in performance. A continuous loop exists between feedback, self-efficacy, and performance.

The majority of research on feedback's effects suggests the seemingly obvious finding that positive feedback increases self-efficacy and resulting performance, and negative feedback usually has a detrimental effect on self-efficacy levels and performance (Krueger & Dickson, 1994; Duke & Henninger, 1998; Nease, Mudgett, & Quiñones, 1999; Brunot, Huguet, & Monteil, 2000; Ayoun, 2001). Furthermore, there are gender differences in response to feedback: Women tend to believe and accept feedback more readily than men, and feel more compelled to act on the feedback they receive, whereas men tend to discount feedback if it disagrees with their own perceptions (Cervone & Palmer, 1990; Pulford & Colman, 1997; Johnson & Helgeson, 2002; Steinmayr & Spinath, 2009). Regardless of gender, recipients of feedback may reject feedback that conflicts significantly with established self-perceptions (Nease, Mudgett, & Quiñones, 1999).

Much of the research indicates that feedback has a potent effect on self-efficacy and performance. Positive feedback can increase self-efficacy for completing tasks. Individuals who believe they can accomplish a specific task are more likely to succeed than those who doubt their abilities. High self-efficacy beliefs in music may help individuals achieve in music. However, self-efficacy alone will not enable individuals to achieve more than their musical aptitude levels allow. It may be interesting to know if feedback can affect scores on a musical aptitude assessment. More would be learned about the impact of feedback and the nature of musical aptitude. There are no studies investigating whether feedback can affect performance on such a test. According to Gordon (1999), musical aptitude stabilizes by age 9 and cannot be increased after this age. Thus, feedback should have no effect on musical aptitude test performance.

Problem and Research Questions

The purpose of the current study was to investigate the effect of positive and negative feedback on musical aptitude test performance between males and females, and music- and non-music majors. The effect of positive and negative feedback on musical aptitude self-efficacy was also tested for both groups of participants. The study addressed these research questions:

1. Were there statistically significant differences between the change scores of participants who have received positive and negative feedback?

2. What was the interaction effect of feedback and gender on these differences?

3. What was the interaction effect of feedback and major on these differences?

4. What was the effect of feedback on participants' self-efficacy for performing on a musical aptitude test?

Population and Sample

Participants were music- and non-music majors recruited from two universities in the Central Southwestern United States. One was a large state university and the other a small private university. Music- and non-music majors were recruited from intact classes at both universities. Data was gathered during regular class meetings at both universities. Class instructors read the recruiting statement to their classes (see Appendix A), with the exception of the non-music major class at the private university. The primary investigator is the instructor for that class, so the recruiting statement was read to the class by the department chair. For all participating classes, the consent forms (see Appendix B) were distributed and collected at the time the recruiting statement was read.

Research Design

Feedback to participants was manipulated, requiring the use of deception

at various times during the study. First, participants were informed that the purpose of the study was to investigate the effect of immediate retesting on a musical aptitude measure. This deception was necessary because test performance may have been affected if participants knew the purpose of the study was to assess the effect of feedback on musical aptitude test performance. Second, participants were led to believe that their pretests would be rated while they completed the questionnaire, when in fact randomly generated ratings were assigned instead. Third, participants were informed that there were three possible ratings for the pretest: high, indicated by a plus sign (+), middle: a check mark (√), and low: a minus sign (-). Actually, only high and low ratings were given. This deception was used because it was reasonable to expect test results to fall into high, middle, and low ratings, though no middle ratings were assigned.

Immediately prior to taking the pretest, participants completed a brief Musical Aptitude Self-Efficacy Scale (See Appendix C), based on a model by Bandura (2006) in which participants answered questions regarding their musical self-efficacy. Participants rated their perceived efficacy to correctly identify a tonal difference between two similar melodies, a rhythmic difference between two similar melodies, or no difference between two melodies. These were the same constructs measured in the pretest. Reliability was high for the MASES; Cronbach's Alpha for the MASES was .949. Gordon's (1989) Advanced Measures of Music Audiation (AMMA) was administered according

to prescribed instructions as a pretest to the participants. The reliability coefficient for the AMMA is .88 for music majors and .83 for non-music majors. Three practice questions were included before the actual test began to acquaint participants with the test format. Test questions consisted of 30 pairs of melodies in which each pair was either exactly alike or differed tonally or rhythmically. Participants indicated their answers by coloring in the appropriate oval on the answer sheet (See Appendix D). The AMMA took approximately 16 minutes to administer. At the end of the pretest, answer sheets were collected and participants completed a brief questionnaire (see Appendix E) in which they indicated their age, gender, and whether they were music majors or non-music majors. Participants also were instructed to write a few sentences addressing their beliefs about their own musical aptitude. Age, gender, and major were the only information actually used from the questionnaire. The section regarding musical aptitude was used merely to occupy participants' time and deceive them into believing the pretests were being rated during that time. The questionnaire was completed in 5-8 minutes.

While participants completed the questionnaire, the principal investigator left the test room and noted pretest ratings of high (+) or low (-) on the backs of the posttest answer sheets using a computer-generated random number sequence in which even numbers represented high pretest scores and odd numbers represented low pretest scores (see Appendix F). It was important to note once again that participants were deceived intentionally. To protect participants'

privacy, posttest answer sheets were handed back individually to participants face-up so that negative and positive feedback (high or low ratings) was not visible. Participants then viewed a PowerPoint® slide (See Appendix G) instructing them to turn their test sheets over to see the feedback rating of a plus sign, a check mark, or minus sign. The slide revealed the meaning of the ratings. Again, it should be noted that no participant received a check mark for "middle" rating.

The AMMA was administered again, this time omitting the recorded instructions and three practice questions. This administration took approximately 13 minutes. The principal investigator and classroom instructor collected the posttests from each participant, receiving them face-up, to protect privacy. Immediately upon completion of the posttest, participants completed the Musical Abilities Assessment Inventory once more. Finally, a debriefing statement (see Appendix H) was distributed explaining the deception used in the study and why it was necessary. Participants were given the opportunity to withdraw their data from the study without penalty. At each of the data collections a random drawing for a $10 iTunes and a $25 Olive Garden gift card was held, and all participants were offered candy bars as thanks for their participation.

Each test was graded manually. Scoring the answer sheet involved 2 processes: computing a raw score by subtracting the number of incorrect

answers from the number of correct answers, and adjusting the raw scores to avoid possible negative scores. The AMMA included four grading masks for manual grading: two masks for the tonal subscale and two masks for the rhythmic subscale. Using the first mask for tonal answers, the correct answers were noted; the second mask was applied to the test sheet to find incorrect answers. As per AMMA scoring procedure, twenty points were added to the first score to avoid the possibility of a negative score. A negative score would result, for instance, if the total number of wrong Tonal answers were higher than the number of right Tonal answers. The second score (incorrect answers) was subtracted from the first (correct answers). This procedure was repeated with the two masks for Rhythm answers. The Tonal score and Rhythm score was added together to make the Total Raw score. This grading process took place for each participant's pretest and posttest. The highest possible Tonal raw score was 40, the highest possible Rhythm raw score was 40, and the highest possible composite score was 80. Maximum scores are rarely attained (Gordon, 1989).

Data Analysis

Although participants were students recruited from both a state and a private university, comparisons between the two groups were not of interest in the current study. Therefore, these groups were considered as one group of "university students" defined by the age range of 18-22 years. Nevertheless, preliminary comparisons between the two groups were conducted using

demographic information and musical aptitude test scores to determine if differences between these groups were a confounding variable on the comparisons of interest. Pretest means for each group are as follows: Private University (M = 51.09, SD = 7.62); State University (M = 55.56, SD = 8.17). Differences in scores were not great enough to consider their difference a confounding variable in the study.

For this quasi-experimental design, a 2x2x2 factorial MANOVA (Feedback x Major x Gender) was performed using the pretest and posttest scores, followed by analyses of main effects and interactions. The dependent variable was the difference between AMMA pretest and posttest scores. Independent variables were gender, feedback, and major. A test of power was run to ensure an appropriate N for each group in the study given the statistical procedures chosen. Post-hoc univariate tests investigated the effect of feedback on AMMA pre- and posttest performance, as well as the effect of feedback on self-efficacy for performing on a musical aptitude test. Results of post-hoc analysis included the Bonferroni/Dunn method to avoid Type I error.

Summary

In the current study, data were collected from intact classes of music majors and non-majors from a state university and a private university. Participants from both universities were combined into one group based on AMMA pretest scores and the narrow age range of 18-22, which unified the

groups into one population for the purposes of this study.

The effect of feedback has been a popular research topic for several decades. There is little doubt that feedback has the power to effect change in performance. There are gender differences in the way feedback is received, and feedback that contradicts established self-efficacy levels is likely to be rejected by both genders. It was interesting to discover whether feedback had a significant effect on males versus females and music majors versus non-music majors using a measure that tested musical aptitude, a stable construct that should not have been affected by feedback.

Chapter 4

Results

The purpose of this study was to investigate the effect of positive and negative feedback on self-efficacy levels regarding musical aptitude. Participants were university students from two Central Southwestern universities-- one state and one private-- in groups according to gender and major: music majors and non-majors. Gordon's (1989) Advanced Measures of Music Audiation (AMMA) was used as both the pretest and the posttest. The AMMA tested the ability to determine whether two short melodies were exactly alike, different tonally or different rhythmically. Through the use of a researcher-designed three-question Musical Aptitude Self-Efficacy Scale (MASES), participants were asked to rate their level of efficacy related to determining whether two short melodies were exactly alike, or contained tonal or rhythmic differences. The MASES was administered twice: once prior to the pretest (MASES 1), and again prior to the posttest (MASES2). Participants were randomly assigned to manipulated positive and negative feedback groups immediately before the administration of the AMMA posttest. The administration of the AMMA posttest began approximately ten minutes after the completion of the AMMA pretest. Participants also completed a questionnaire requesting gender, age, and college major between the administration of the AMMA pretest and posttest. Although participants reported significant changes in their self-efficacy levels after receiving

feedback, there were no statistically significant changes in AMMA posttest scores.

The research questions were these:

1. Would there be statistically significant differences between the AMMA pre-test/posttest scores of participants who received positive and negative feedback?

2. What was the interaction effect of feedback and gender on these differences?

3. What was the interaction effect of feedback and major on these differences?

4. What was the effect of feedback on participants' self-efficacy for performing on a musical aptitude test?

Data Analysis

The results reported in this chapter are arranged according to research question. Descriptive statistics for the study were compiled; means and standard deviations may be found following each research question. The research design used college major, gender, and feedback as independent variables. Dependent variables were two scores: the difference between musical aptitude self-efficacy scores (MASES 1 and MASES 2) and difference between pretest and posttest scores from Gordon's (1989) Advanced Measures of Music Audiation (AMMA). Statistics were computed using the Statistical Package for the Social Sciences (SPSS) version 20 software. A one-way multivariate analysis of vari-

ance (MANOVA) was conducted to investigate the interaction effect of positive and negative feedback on males and females, music majors and non-music majors. Post hoc analyses of variance (ANOVAs) were conducted as follow-up tests on dependent variables AMMA pretest/posttest scores, and MASES scores to determine further effects of feedback on gender and major.

Distribution Figures

Distribution figures can be seen in Table 1. To determine that students from both the private and state university could be combined as one group of participants, means and standard deviations from the AMMA pretest were compared between participants from both schools. They are as follows: Private University (M = 51.09, SD = 7.62) and State University (M = 55. 56, SD = 8.17). Music majors from both universities each were combined as one group, as were non-music majors, on the same basis.

Table 1

Distribution by School, Gender, and Major

	N	% of Total
Private University	131	59
State University	91	41
Males	116	52
Females	106	48
Music Majors	70	32
Non-Music Majors	152	68

A slight majority of participants, 59%, were from the private university. The number of male and female participants was almost equal, with males outnumbering females. Non-music majors outnumbered music majors, comprising 68% of participants. The total number of participants was 222.

MANOVA Results

A factorial multivariate analysis of variance was conducted to determine the effect of feedback on AMMA pretest and posttest score differences and MASES 1 and 2 score differences by gender and major. As the MANOVA was the main statistical procedure, its results will be discussed first in Table 2, followed by the Tests of Between-Subjects Effects in Table 3. Using the Bonferroni method, the MANOVA was tested at the $p < .025$ level to avoid Type I error.

Table 2

The Effect of Feedback on AMMA Pre/Posttest and MASES 1 and 2 Score Differences by Gender and Major MANOVA

	Λ	F	df1	df2	Sig
.					
Feedback	.767	32.88	2.00	216	.000*
Feedback x Gender	.989	1.00	2.00	216	.294
Feedback x Major	.984	1.71	2.00	216	.182
Feedback x Gender x Major	.999	0.130	2.00	216	.878

*Significant at the .025 level

Feedback had a significant effect ($p < .0001$). None of the combined interactions of the AMMA pretest/posttest and the MASES 1 and 2 showed a significant effect. Feedback did not have a significant effect by gender, at $p = .294$. At $p = .878$, the interaction of feedback, gender, and major was the least significant result in the MANOVA.

Results indicate no significant differences in AMMA pretest/posttest scores for gender or major. However, a significant effect of feedback was found on MASES 1/MASES 2 score differences ($p < .0001$) for all participants.

Research Question #1: Would there be statistically significant differences between the AMMA pretest/posttest scores of participants who received positive and negative feedback?

The results of the multivariate analysis of variance (MANOVA) showing effects of feedback on major and gender, both between subjects and within subjects may be seen in Table 3.

Descriptive statistics are shown for participants who received positive feedback versus participants who received negative feedback. Means and standard deviations are shown for AMMA Pretest scores, before feedback was given, and for AMMA Posttest scores, after feedback was given to participants. Results may be found in Table 4

Table 3

Tests of Between-Subjects Effects of AMMA Pretest/Posttest Differences and MASES 1 and 2 Differences MANOVA

	Dependent Variable	df	Mean Square	F	Sig.
Major	MASES1/2 diff.	1	2008.32	3.04	.083
	AMMA pre-post diff.	1	12.20	0.206	.650
Gender	MASES1/2 diff.	1	1592.12	2.41	.122
	AMMA pre-post diff.	1	10.19	0.172	.679
Feedback	MASES1/2 diff.	1	41971.89	63.59	.000*
	AMMA pre-post diff.	1	31.39	0.530	.467
Major/Gender	MASES1/2 diff.	1	119.51	0.181	.671
	AMMA pre-post diff.	1	3.30	0.081	.814
Major/Feedback	MASES1/2 diff.	1	1733.36	2.64	.105
	AMMA pre-post diff.	1	0.092	0.002	.969
Gender/Feedback	MASES1/2 diff.	1	915.22	1.39	.239
	AMMA pre-post diff.	1	50.81	0.85	.358
Major/Gender/FB	MASES1/2 diff.	1	931.88	1.42	.234
	AMMA pre-post diff.	1	28.60	0.478	.490

* Significant at the .025 level

Table 4

Means and Standard Deviations for Effect of Feedback on AMMA Pretest and Posttest Scores

		N	Mean	SD
AMMA Pretest Total	Positive Feedback	105	53.48	7.83
	Negative Feedback	117	52.75	8.40
	Total	222	53.09	8.13
AMMA Posttest Total	Positive Feedback	105	56.60	8.78
	Negative Feedback	117	55.09	9.26
	Total	222	55.81	9.05

The means for the participants who received positive feedback showed an increase, but not a statistically significant increase. Likewise, the means for participants who received negative feedback also increased from pretest to posttest. Standard deviations were very similar for participants pretest to posttest who received positive or negative feedback. For the pretest, the standard deviation was *SD* = 7.83 for participants who received positive feedback versus *SD* = 8.4 for participants who received negative feedback. On the posttest, the standard deviation for participants in the positive feedback treatment was *SD* = 8.78 versus the standard deviation for those in the negative feedback treatment, at *SD* = 9.05. The collective standard deviation was slightly higher for the posttest (pretest *SD* = 8.13 versus posttest *SD* = 9.05) indicating a greater variance in AMMA posttest scores after feedback.

A post hoc ANOVA was conducted to investigate the effect of feedback on the difference between the AMMA pretest and posttest scores between and within groups who received positive and negative feedback. Results were tested using a Bonferroni adjustment at $p < .025$ to avoid Type I error. Table 5 displays the ANOVA results.

Table 5

The Effect of Feedback n AMMA Pretest/Posttest Score Differences ANOVA

		df	Mean Square	F	Sig.
Pretest	Between Groups	1	29.28	0.449	
	Within Groups	221	66.26	0.503	
	Total	222			
Posttest	Between Groups	1	127.09	1.55	
	Within Groups	221	81.75	0.214	
	Total	222			

*Significance at .025 level

Results indicated no significant effect of feedback on AMMA Posttest scores (p = .214).

Research Question #2: What was the interaction effect of feedback and gender on these differences?

Table 6 features means and standard deviations of AMMA pretest and posttest scores for males and females.

Table 6

Means and Standard Deviations of Pretest and Posttest AMMA Scores for Gender

	AMMA Pretest		AMMA Posttest	
Gender	M	SD	M	SD
Females	53.14	7.77	55.53	9.08
Males	53.06	8.47	56.06	9.05

As Table 6 illustrates, pretest scores for males and females were almost identical. Males showed more variability in the standard deviation than females, but the difference was negligible. Posttest scores were higher for both groups, with males scoring slightly higher on the posttest than females. Thus, positive and negative feedback had no significant effect on AMMA posttest scores for males or females. The standard deviation was higher for both genders on the posttest than on the pretest, indicating a greater variance in scores than on the pretest. This variance may suggest an effect of feedback and test effect for both genders.

The means for males and females who received positive feedback versus males and females who received negative feedback may be seen in Ta

Table 7

Means and Standard Deviations for Males and Females Receiving Positive and Negative Feedback

Feedback	Gender	Mean	Standard Deviation
Positive FB	Male	-3.75	6.68
	Female	-2.37	7.19
	Total	-3.12	6.92
Negative FB	Male	-2.30	8.10
	Female	-2.41	8.47
	Total	-2.35	8.25
Total	Male	-3.01	7.44
	Female	-2.39	7.88
	Total	-2.72	7.64

AMMA posttest mean differences were non-significant for gender.

A one-way Analysis of Variance (ANOVA) was conducted as a post hoc test to the MANOVA to investigate the interaction effect of feedback and gender on AMMA posttest scores. Results can be seen in Table 8.

Table 8

Interaction Effect of Feedback and Gender on AMMA Posttest Scores ANOVA

	df	Mean Square	F	Sig.
Feedback	1	27.43	0.465	0.496
Gender	1	22.25	0.378	0.540
Feedback x Gender	1	30.51	0.518	0.473

No statistically significant effect was found for the interaction of positive and negative feedback and gender on AMMA posttest scores.

Research question #3: What was the interaction effect of feedback and major on these differences?

Descriptive statistics for the effect of positive and negative feedback on the AMMA pretest/posttest score differences for music majors and non-music majors may be found in Table 9. A post-hoc one-way ANOVA follows in Table 10.

Table 9

Means and Standard Deviations for the Effect of Feedback on AMMA Posttest Scores by Major

Major	Feedback	Mean	Std. Deviation
Music Major	Positive FB	-2.67	5.87
	Negative FB	-2.05	8.10
	Total	-2.35	7.06
Non-Music Major	Positive FB	-3.33	7.40
	Negative FB	-2.49	8.36
	Total	-2.89	7.91
Total	Positive FB	-3.12	6.92
	Negative FB	-2.36	8.25
	Total	-2.72	7.64

When examining AMMA posttest scores of music majors, it was found that the scores of those receiving positive feedback were .62 points higher than scores of those receiving negative feedback. Means scores for non-music majors who received positive feedback were .84 points higher than for non-music majors who received negative feedback. Means for all participants combined were .76 points higher for those who received positive feedback over those who received negative feedback.

Table 10

Interaction Effects of Feedback and Major on AMMA Posttest Scores ANOVA

	df	Mean Square	F	Sig.
Music Major	1	14.46	.245	.621
Feedback	1	25.67	.435	.510
Music Major x Feedback	1	0.596	.010	.920

*Significance at the 0.025 level

No significant interaction effect was found of feedback and major on AMMA posttest scores.

Research Question #4: What was the effect of feedback on participants' self-efficacy for performing on a musical aptitude test?

The previous three research questions dealt with the effect of feedback, gender, and major on AMMA posttest performance. Question 4, however, was concerned with participants' self-efficacy for performance on the AMMA. Participants reported their self-efficacy levels from 0 to 300 on the researcher-designed Musical Aptitude Self-Efficacy Scale (MASES). Table 11 shows the descriptive statistics for MASES 1/MASES 2 score differences by gender.

Table 12 is an ANOVA showing between-subjects effect of feedback on gender as indicated by MASES 1/MASES2 score differences. Using the Bonferroni method, testing was conducted at the .025 level of significance to avoid Type I error.

Table11

Descriptive Statistics for MASES1/MASES 2 Score Differences by Gender

Gender	Feedback	Mean	Std. Deviation
Male	Positive FB	-1.70	29.47
	Negative FB	23.39	26.06
	Total	11.06	30.40
Female	Positive FB	-9.08	22.44
	Negative FB	20.93	24.15
	Total	7.34	27.7
Total	Positive FB	-5.07	26.62
	Negative FB	22.17	25.06
	Total	9.28	29.14

Mean score differences between males and females indicated that females were more influenced by positive feedback than males. After receiving positive feedback, MASES2 mean scores for females increased 7.38 points higher than mean scores for males. After negative feedback, MASES 2 mean scores for females were 2.46 points lower than mean scores for males. Overall, for both genders there was an effect of positive and negative feedback as reflected by MASES 2 mean scores, indicating that feedback affected participants' self-efficacy regarding their musical aptitude.

There was a significant effect ($p < .0001$) of feedback on MASES1/MASES 2 score differences. MASES 2 scores increased for participants who received positive feedback and MASES 2 scores decreased for partic-

ipants who received negative feedback. The interaction between feedback and gender indicated no statistical significance.

Table 12

The Interaction Effect of Feedback and Gender on MASES 1/MASES 2 Score Differences ANOVA

	df	Mean Square	F	Sig
Gender	1	1334.381	2.0	.158
Feedback	1	41845.95	62.9	.000*
Gender/Feedback	1	333.34	0.502	.479

*Significant at the .025 level

The interaction effect of feedback and major on the MASES 1/MASES 2 score difference may be seen in Table 13, which shows descriptive statistics.

Total mean differences increased on the MASES 2 for all participants who received positive feedback. The MASES 2 score was subtracted from the MASES 1 score to find scores differences. A negative score indicated a higher MASES 2 score than MASES 1 score. However, after receiving positive feedback, means for non-music majors increased more than means for music majors. Among participants who received negative feedback, MASES 1/MASES 2 mean differences for music majors and non-majors was almost identical. The overall effect of feedback on major indicated a greater change in MASES 1/MASES 2 scores for non-music majors than for music majors.

Table 13

Descriptive Statistics: The Effect of Feedback on MASES 1/MASES 2 Score Differences by Major

Feedback	Major	Mean	Std Deviation
Positive FB	Music Major	3.20	24.72
	Non-Music Major	-9.04	26.76
	Total	-5.07	26.62
Negative FB	Music Major	22.44	21.80
	Non-Music Major	22.04	26.50
	Total	22.17	25.06
Total	Music Major	13.10	25.04
	Non-Music Major	7.52	30.76
	Total	9.28	29.14

Table 14 shows the ANOVA results for the effect of feedback and major on MASES 1/MASES 2 score differences. Results were tested using the Bonferroni method for significance at $p < .025$ to avoid a Type I error.

A significant effect was found for feedback on all MASES 1/MASES 2 score differences. The interaction of feedback and major was not significant ($p = .111$). Thus, positive and negative feedback had a highly significant effect on participants, with a greater effect on music majors.

Table 14

Interaction Effect of Feedback and Major on MASES 1/MASES 2 Score Differences ANOVA

	df	Mean Square	F	Sig
Feedback	1	30293.43	46.13	.000*
Music Major	1	1911.63	2.91	.083
Feedback/Music Major	1	1680.17	2.59	.111

*Significant at the 0.025 level

Conclusion

Feedback did not have a significant effect on AMMA posttest performance. Furthermore, neither the interaction effect of feedback and gender, nor feedback and major, had a statistically significant effect on AMMA posttest performance. However, there was significant effect ($p < .0001$) of feedback on self-efficacy levels, as reflected by the MASES 1 and MASES 2 score differences. Though not significant, there was a greater effect on females than males and a greater effect on music majors than on non-music majors. Self-efficacy levels, as reflected by MASES 2 scores, were affected by feedback, though the change in self-efficacy did not affect performance on the AMMA posttest.

Chapter 5

Findings, Conclusions, and Implications

The purpose of the study was to investigate the effect of feedback on musical self-efficacy and performance on a musical aptitude test. There were four research questions: 1) Are there statistically significant differences between the Advanced Measures of Music Audiation (AMMA) pretest/posttest scores of participants who received positive and negative feedback? 2) What is the interactive effect of feedback and gender on AMMA pretest/posttest score differences? 3) What is the interactive effect of feedback and major on AMMA pretest/posttest score differences? 4) What is the effect of feedback on participants' self-efficacy for performing on a musical aptitude test?

Participants were 222 music majors and non-majors from a central Southwestern state and a central Southwestern private university. Music majors were tested during a regular music class period; non-majors were tested during a music appreciation class meeting. Participants were administered Gordon's (1989) Advanced Measures of Music Audiation (AMMA), a standardized musical aptitude test designed for high school and university students, as a pretest and a posttest. The test items were short pairs of melodies that were either exactly alike, or different tonally or different rhythmically.

To measure participants' task-specific self-efficacy, a Musical Aptitude Self-Efficacy Scale (MASES) was administered twice to participants: first, before any AMMA testing began, and second, immediately after receiving false feedback about their AMMA pretest scores. Reliability for the MASES was high: Cronbach's Alpha for MASES 1 was .949, and for MASES 2 was .963. The first MASES administration was to establish participants' musical aptitude self-efficacy before manipulation occurred, and the second MASES administration was to learn if manipulated feedback had an effect on participants' musical aptitude self-efficacy. Designed by the researcher with assistance from a committee member, the MASES was patterned after the model Bandura (2006) suggested as most effective for measuring self-efficacy levels. Based on the 0-100 continuum recommended by Bandura, the scale provided more exact measurement of self-efficacy levels than scales with narrower parameters. The three questions on the MASES addressed participants' self-efficacy to answer the questions found on the AMMA: whether two melodies were exactly alike, different tonally, or different rhythmically. In both administrations of the MASES, participants submitted three scores between 0 and 100 to indicate their self-efficacy level for each of the three types of melodies tested by the AMMA.

After the AMMA pretest, participants answered a short demographic questionnaire while the pretests were marked with false ratings. The test scores were randomly manipulated by the researcher so that approximately half the participants received positive feedback and half received negative feedback in the

form of a positive or negative rating on the back of their blank posttest. Participants were shown a PowerPoint® slide explaining that those who received a plus sign scored very high on the AMMA pretest, and those receiving a minus sign scored very low on the AMMA pretest. Immediately after receiving manipulated feedback, participants completed a second MASES indicating their level of self-efficacy about musical aptitude. Finally, the AMMA was administered again, this time as a posttest. At the completion of the AMMA posttest, the deception was explained to participants and then they were dismissed. The entire data collection took place within a 50-minute class period.

Research Findings

Statistical measures included a multiple analysis of variance (MANOVA) and post hoc analyses of variance (ANOVAs). Independent variables were major, gender, and feedback, and dependent variables were pretest/posttest score differences and MASES 1/MASES2 score differences. Significant main effects were not found for any of the research questions except for the effect of feedback on participants' musical self-efficacy ($p < .0001$) both for music majors and non-music majors.

Research Question 1: Will the difference between the AMMA pretest/posttest scores be statistically significant among participants who receive positive and negative feedback?

When considering the AMMA pretest and posttest scores, there was a slight increase in posttest scores, regardless of the type of feedback received (AMMA pretest $M = 53.09$, AMMA posttest $M = 55.81$). This increase was not significant ($p = .214$). The AMMA pretest and posttest were administered twice within a fifty-minute period. Although the AMMA measured musical aptitude, a stabilized construct among university-age students, test effect may have been a factor in that participants were better able to answer more questions correctly upon hearing the same questions again immediately after the pretest. Regardless of positive or negative feedback, participants may have remembered enough of the pretest questions to improve their performance on the posttest. Thus, the posttest in the present study contained some elements of an achievement test, as the scores went up on the posttest, whether participants received positive or negative feedback, though scores remained within the norms of the AMMA reliability.

The results of the present study do not confirm Bridgeman's (1974) study in which seventh-graders were given false feedback from a scholastic aptitude test immediately before taking a portion of the Lorge-Thorndike Intelligence Test. In Bridgeman's study, participants receiving false positive feedback had significantly higher scores ($p < .01$) than those receiving false negative feedback. In the present study, performance on the AMMA posttest was not significantly affected by false positive or negative feedback. While both the Lorge-Thorndike Intelligence Test and the AMMA measured cognitive con-

structs, the two constructs differ in that intelligence is fairly malleable whereas musical aptitude is stabilized around age 9 (Gordon, 1989). The age difference in each group of participants may also be a factor. It is possible that nine-year-olds are more receptive to manipulative feedback than college-age individuals. Thus, musical aptitude scores would be less likely to change than scores on an intelligence test. A change in musical self-efficacy, even a significant change such as in the present study, did not affect performance on the AMMA posttest.

Research Question 2: What is the interaction effect between feedback and gender on AMMA posttest score differences?

Literature on feedback and gender suggests that women and men react differently to positive and negative feedback. When confronted with negative feedback, men exhibited more optimism in their abilities than women, and tended to be overconfident (Pulford & Colman, 1997; Seligman, Nolen-Hoeksema, Thornton, & Thornton, 1990; Cervone & Palmer, 1990). Moreover, women's reaction to feedback tended to be more polarized than men's (Beyer & Langenfeld, 2000). Though scores on the whole increased for both genders from AMMA pretest to posttest, there were no significant differences in scores of the AMMA pretest/posttest between males and females (p = .294), suggesting that the reaction to positive and negative feedback was similar in males and females.

In comparing the AMMA pretest/posttest scores between genders in the present study and the researcher's previous study (Hutton, 2006), the effect of feed-

back on gender was not significant in either study (Previous study $p = .20$; Current study $p = .29$). In Hutton's previous study, males' AMMA pretest scores were higher than females' AMMA pretest scores (Males $M = 51.54$, $SD = 19.17$ versus Females $M = 46.63$, $SD = 19.95$ out of a perfect score of 80). However, the females' posttest scores were higher than the males' scores. Not only did the females' scores increase from pretest to posttest, but their scores were also higher than the males' scores on the posttest. The increase in females' scores could be attributed to their reaction to positive feedback. Males' scores increased very little, regardless of direction of the feedback. In the current study, there was little difference in AMMA pretest scores between males and females (Males $M = 53.06$, $SD = 8.47$ versus Females $M = 53.14$, $SD = 7.77$). AMMA posttest scores between genders likewise were similar in the current study (Males $M = 56.06$, $SD = 9.05$ versus Females $M = 55.53$, $SD = 9.08$). In the current study, the males' AMMA posttest scores were 0.53 higher than the females' scores. Scores increased from pretest to posttest for both males and females, but the increase was not significant for either gender.

The main difference between the testing circumstances in Hutton's (2006) study and the present study is the span of time between the administration of the pretest and posttest. In the 2006 study, there were two days between test administrations, whereas in the present study, both the pretest and posttest were taken in a single 50-minute period of time. It was possible that the two-day gap in test administrations allowed participants to forget much of the AMMA pretest, sub-

sequently feeling more vulnerable to the effects of manipulated feedback. That the feedback had a significantly greater effect on females than males is in keeping with research findings (Pulford & Colman, 1997; Seligman, Nolen-Hoeksema, Thornton, & Thornton, 1990; Cervone & Palmer, 1990). However, feedback did not have a significant effect on AMMA posttest scores of males or females in the present study. It is interesting to note that the scores on the AMMA pretest and posttest were higher in the current study than in the 2006 study. This is likely because music majors were included in the current study and only non-music majors participated in the 2006 study.

Research Question 3: What is the interaction effect between feedback and major on AMMA posttest score differences?

In the present study, the AMMA pretest-posttest score differences of music majors and non-music majors did not change significantly. Scores were higher on the AMMA posttest for both groups (Majors M = -2.35, SD = 7.06; Non-Music Majors M = -.2.89, SD = 7.91). The negative scores indicated that the posttest scores were higher than the pretest scores for both groups. Although feedback affected self-efficacy levels for both music majors and non-music majors, feedback did not have a significant effect on the AMMA posttest for either group (p = .920). This could have been due, in part, to the AMMA's validity in measuring the specific constructs of tonal recognition and tonal memory. Additionally, participants' musical aptitude had been long established, having stabi-

lized around age 9 (Gordon, 1989). Thus, the combination of the AMMA's high validity and reliability and participants' well-established musical aptitude may have provided substantial protection from the effect of feedback to influence AMMA posttest performance.

Research Question 4: What is the effect of feedback on participants' self-efficacy?

The researcher-designed Musical Aptitude Self-Efficacy Scale (MASES) was used to measure self-efficacy levels specific to musical aptitude. The MANOVA results revealed a significant main effect with the MASES 1/MASES 2 score differences. There was a significant effect of feedback ($p < .0001$) on participants' self-efficacy levels in the study as indicated by the MASES 1 and 2 score differences. Though not significant at $p < .05$, results of a post hoc ANOVA suggest that music majors' self-efficacy levels were more affected by feedback than non-majors' self-efficacy ($p < .083$). An effect of feedback on music majors' self-efficacy was expected, as music majors would likely have placed a higher importance in their musical identity than non-majors.

This result mirrors a study involving the effect of false feedback on career identity by Brunstein and Gollwitzer (1996). In this study students were administered a concentration game and then given false failure feedback. The false failure feedback had detrimental effects ($p < .001$) on the self-efficacy of medical students and in turn had a significant effect on the medical students' ability

to perform on a second task involving mental concentration because the failure feedback reflected potential professional shortcomings. Hence, it follows that manipulated failure and success feedback regarding musical aptitude would have a more pronounced effect on music majors than it would on non-music majors. Music majors who plan to make music their career would place high importance on having high musical aptitude.

On the MASES 1 and 2 participants reported their self-efficacy to perform well on the AMMA pretest-posttest. The AMMA pretest-posttest involved musical cognitive reasoning, with a number of the AMMA test items being very difficult. In a study investigating self-efficacy and cognitive challenges, Stone (1994) suggested that individuals initially expressed over-confidence in their abilities when faced with cognitive challenges. This may have occurred with the MASES 1, in which participants indicated their level of self-efficacy to detect tonal and rhythmic changes in melodies. Scores on the MASES 2 were lower after negative feedback, but a portion of the MASES 2 scores were lower after positive feedback, as well. This could be due to several factors, among them test fatigue from two administrations of the AMMA within a 50-minute period. Several students remarked that they were exhausted after two administrations of the AMMA in close succession because of the intense concentration required to take the test. Additionally, after taking the AMMA as a pretest and noting the difficulty of several of the test items, some participants may have had lower self-efficacy in their ability to answer test items correctly on the AMMA posttest.

Thompson and Richardson (2001) suggested that test anxiety might interfere with performance. While test anxiety may or may not have influenced performance on the AMMA pretest, it may have been a factor in the lower self-efficacy levels in performing on the AMMA posttest, possibly after participants experienced the difficulty of the test, and/or received false negative feedback.

Research into feedback's effect on gender indicates that males and females react differently to feedback (Pulford & Colman, 1997; Seligman, Nolen-Hoekstra, Thornton & Thornton, 1990; Cervone & Palmer, 1990). Thus, in the current study it was expected that there would be a gender difference in reaction to feedback, with women exhibiting more reaction to feedback than males. MASES1 scores were almost identical for males and females indicating very similar self-efficacy levels regarding musical aptitude prior to feedback (Males M =160.9 out of a possible 300; SD = 82.73 versus Females M = 159.1, SD= 84.17). After feedback, however, a post-hoc ANOVA indicated a significant main effect of feedback on self-efficacy for both males and females ($p < .0001$). While there was not a significant difference between genders ($p = .158$) in MASES 1/MASES2 scores, males rated their self-efficacy higher after both positive and negative feedback than females did. Males' mean scores were 7.97 points higher than females' scores, regardless of positive or negative feedback, indicating that, while the difference is not statistically significant, males sustained higher overall self-efficacy levels than females, whether they received positive or negative feedback. It is interesting to note, however, that among par-

ticipants that received positive feedback, females had higher MASES 2 scores than males (Females *M* = -9.08, *SD* = 27.7 versus Males *M* = -1.70, *SD* = 29.47). Although this difference is not statistically significant, it does support the findings that females may accept feedback differently than men. Specifically, females may be more receptive to positive feedback than men.

Limitations of the Study

The primary limitation in the current study was the time restriction for testing. Because participants were university students, data were collected when students were available. Regular class meetings were chosen to allow for the largest number of participants. Data collection took place during one 50-minute class session in order to cause the least interruption of the regular class routine for students and the professors who taught them. In the previous study (Hutton, 2006), data collection took place two days apart; the AMMA pretest was administered on a Tuesday and the posttest given the following Thursday. In the current study, the time limitations may have contributed to possible test effect, allowing posttest scores to be higher. It is possible that participants remembered enough of the material from the pretest to perform more successfully on the posttest, regardless of the direction of the feedback.

Another limitation of the study was the disproportion of music majors to non-music majors as participants. Combining the music majors from both universities produced 70 music majors versus 152 non-music majors. Non-music

major participants were students from three music appreciation classes taught exclusively for non-music majors. A state music convention was occurring at the same time as data collection, in which several music majors were involved, that further reduced the numbers of potential music majors who could have participated.

Conclusions and Implications

It is interesting to note the dichotomy of results for the AMMA pretest-posttest score differences and the MASES1/MASES 2 score differences. The false feedback produced a significant main effect ($p < .0001$) on self-efficacy levels, as indicated by the MASES1 and 2 scores, yet post hoc tests indicated that feedback did not significantly affect performance on the AMMA posttest. As discussed earlier in the chapter, this may be due to several factors, among them the testing situation itself. Practicality required that participants be administered all tests in one class period. Because of the limited time, there was the likely possibility of test effect when participants were administered the AMMA pretest and posttest in quick succession. Although the AMMA is an aptitude test and not an achievement measure, it was not designed for multiple administrations in quick succession. Many of the AMMA posttest scores were higher than the pretest scores, regardless of positive or negative feedback. Thus, some level of achievement may have occurred, as participants remembered parts of the pretest as they took the same test again a few minutes later as a posttest. More re-

search could include an investigation of whether scheduling test administrations days apart would produce different results.

Further research is needed to determine why a significant effect of feedback on AMMA scores was obtained when testing occurred two days apart and no significance occurred when testing took place within a 50-minute class period. The Hutton 2006 study used only non-music majors as participants. Further research is needed to investigate if a greater time difference would have an impact on the effect of feedback on the administration of the pretest and posttest to music majors. Both Hutton's previous study and the current study were conducted using university students as participants. Young adults at this life stage are making important decisions about their careers, regardless of their major. Music majors naturally question whether their skills and musical aptitude are sufficient to sustain a career in music. Research should be expanded to include music professionals, both performers and educators—individuals who have achieved careers in music-- to investigate whether manipulated feedback would affect their self-efficacy.

An interesting dichotomy emerged from this study. On one hand, feedback had a significant effect on self-efficacy levels for musical aptitude. On the other, musical aptitude test performance was not affected. Scores on the AMMA posttest increased, regardless of feedback, which indicates that learning took place from the first test to the second. Participants remembered enough of

the test material to improve their scores on the posttest. For participants who received negative feedback to improve on the second test speaks to the resiliency of the human mind and the robust nature of musical aptitude.

Receiving negative feedback is part of the learning process. Teachers may hesitate to offer negative feedback for fear that it will damage their students. The current study indicates that students are strong enough to not only receive negative feedback, but to improve their test performance after receiving it. This points to the resilient nature of students.

The feedback in the study was in the form of one rating symbol—a plus sign or a minus sign—and a written statement as to the meaning of the rating symbol. No verbal feedback was given to participants. If feedback in the form of one written rating can have a significant effect ($p < .0001$) on the self-efficacy of participants, it causes one to consider how much more impact verbal feedback must have. Verbal feedback is filled with far more nuance and inflection than written communication. Facial expression and body language enhance the message being delivered. It is the privilege and responsibility of music educators to give students feedback in a sensitive and encouraging manner because students are all too willing to believe them. Music educators should take every opportunity to give feedback that is sincere and constructive. The influence of a teacher who treats students with enthusiasm for music learning and a respect for their individuality will have an impact on generations to com

References

Allen, J., & Howe, B. (1998). Player ability, coach feedback, and female adolescent athletes' perceived competence and satisfaction. *Journal of Sport and Exercise Psychology, 20,* 280-299.

Amorose, A., & Weiss, M. (1998). Coaching feedback as a source of information about perceptions of ability: A developmental examination. *Journal of Sport and Exercise Psychology, 20,* 395-420.

Austin, J., & Vispoel, W. (1992). Motivation after failure in school music performance classes: The facilitative effects of strategy attributions. *Bulletin of the Council for Music Research in Music Education, 111,* 1-23.

Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passe compose and imparfait. *The Modern Language Journal, 85,* 226-243.

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

Bandura, A. (1982a). Mechanism in Human Agency. *American Psychologist, 37,* 122-147.

Bandura, A. (1986). *Social Foundations of Thought and Action.* Englewood Cliffs, New Jersey: Prentice Hall.

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology, 25,* 729-735.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28,* 117-148.

Bandura, A. (1997). *Self-Efficacy: The exercise of control.* New York: W.H. Freeman and Company.

Bandura, A. (2006). Guide for creating self-efficacy scales. In Pajares, F. & Urdan, T. (Ed.) *Self-efficacy beliefs of adolescents* (pp. 307-338). Greenwich, Conn.: Information Age Publishing.

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45,* 1017-1028.

Bannister, B. (1986). Performance outcome feedback and attributional feedback: Interactive effects on recipient responses. *Journal of Applied Psychology, 71*, 203-210.

Baron, R. (1988). Negative effects of destructive criticism: impact on conflict, self-efficacy, and task performance. *Journal of Applied Psychology, 73,* 199-207.

Baron, R. (1990). Countering the effects of destructive criticism: The relative efficacy of four interventions. *Journal of Applied Psychology, 75*, 235-245.

Baumeister, R., Hutton, D., & Cairns, J. (1990). Negative effects of praise on skilled performance. *Basic and Applied Psychology, 11*, 131-148.

Beyer, S. (1998). Gender differences in causal attributions by college students of performance on course examinations. *Current Psychology, 17*, 346-359.

Beyer, S., & Langenfeld, K. (2000). Gender differences in the recall of performance feedback. *Annual Conference of the Midwestern Psychological Association* , (pp. 1-6). Chicago, Illinois.

Black, S., & Weiss, M. (1992). The relationship among perceived coaching behaviors, perceptions of ability, and motivation in competitive age-group swimmers. *Journal of Sport and Exercise Psychology, 14*, 309-325.

Bong, M. (2006). Asking the right questions: how confident are you that you could successfully perform these tasks? In Pajares, F. & Urdan, T. (Eds.) *Self-efficacy beliefs of adolescents* (pp. 287-306). Greenwich, Conn.: Information Age Publishing.

Bong, M., & Clark, R. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychology, 34(3),* 139-153.

Bong, M., & Skaalvik, E. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*, 1-41.

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology, vol. 130(3)*, 353-363.

Bridgeman, B. (1974). Effects of test score feedback on immediately subsequent test performance. *Journal of Educational Psychology, Vol. 66,* 62-66.

Brunot, S., Huguet, P., & Monteil, J. (2000). Performance feedback and self-focused attention in the classroom: When past and present interact. *Social Psychology of Education, 3*, 277-293.

Brunstein, J., & Gollwitzer, P. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology, 70*, 395-407.

Brunstein, J., & Olbrich, E. (1985). Personal helplessness and action control: Analysis of achievement-related cognitions, self-assessments, and performance. *Journal of Personality and Social Psychology, 48*, 1540-1551.

Burnett, P. (1996). Children's self-talk and significant others' positive and negative statements. *Educational Psychology,16(1)* 57-67.

Burnett, P. (1999). Children's negative self-talk and academic self-concepts. *Educational Psychology in Practice, 15(3),* 195-200.

Butler, D., & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65 (3),* 245-261.

Cassady, J., & Johnson, R. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*, 270-295.

Cervone, D., & Palmer, B. (1990). Anchoring biases and the perseverance of self-efficacy beliefs. *Cognitive Therapy and Research, 14(4),* 401-416.

Chapell, M., Blanding, B., Silverstein, M., Takahashi, M., Newman, B., Gubi, A., & McCann, M. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, Vol. 97(2),* 268-274.

Chen, G., Gully, S., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods, 4*, 62-83.

Duke, R., & Henninger, J. (1998). Effects of verbal corrections on student attitude and performance. *Journal of Research in Music Education, 46(4),* 482-495.

Duke, R., & Madsen, C. (1991). Proactive versus reactive teaching: Focusing observation on specific aspects of instruction. *Bulletin of the Council for Research in Music Education, 108*, 1-14.

Dykeman, B. (1994). The effects of motivational orientation, self-efficacy, and feedback condition on test anxiety. *Journal of Instructional Psychology, 21(2),*114-117.

Gist, M., & Mitchell, T. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *The Academy of Management Review, 17(2),*183-211.

Gordon, E. (1989). *Manual for the Advanced Measures of Music Audiation.* Chicago: G.I.A. Publications.

Gordon, E. (1965). *Musical Aptitude Profile*. Chicago: Houghton Mifflin/Riverside Publishing.

Gordon, E. (1999). All about audiation and music aptitudes. *Music Educators Journal, 86(2),* 41-44.

Hallam. S., & Shaw, J. (2002). Constructions of musical ability. *Bulletin of the Council for Research in Music Education, 153(4*), 102-108.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Journal Research, 77(1),* 81-112.

Hong, E. (1999). Test anxiety: Perceived test difficulty and test performance: Temporal patterns of their effects. *Learning and Individual Differences, 11(4), 431-450.*

Hutton, P. (2006). A study of positive and negative suggestion on music aptitude test scores *(unpublished study).*

Jain, S., Bruce, M., Stellern, J., & Srivastava, N. (2007). Self-efficacy as function of attributional feedback. *Journal of School Counseling, 5(4),* 3-22.

Johnson, M., & Helgeson, V. (2002). Sex differences in response to evaluative feedback: A field study. *Psychology of Women Quarterly, 26,* 242-251.

Karl, K., O'Leary-Kelly, A., & Martocchio, J. (1993). The impact of feedback and self-efficacy on performance in training. *Journal of Organizational Behavior, 14 (4),* 379-394.

Korsgaard, A. (1996). The impact of self-appraisals on reactions to feedback from others: The role of self-enhancement and self-consistency concerns. *Journal of Organizational Behavior, 17(4),* 301-311.

Krueger, N., & Dickson, P. (1994). How believing in ourselves increases risk taking: Perceived self-efficacy and opportunity recognition. *Decision Sciences, 25(3),* 385-400.

Lee, J. (1999). Test anxiety and working memory. *Journal of Experimental Education, 67(3),* 218-241.

Margolis, H. & McCabe, P. (2004). Self-Efficacy: A key to improving motivation of struggling learners. *The Clearing House, 77,* 241-249.

McCrystal, R. (1995). *A validity study of the Advanced Measures of Music Audiation among undergraduate college music majors (Doctoral dissertation).* Retrieved from Proquest Dissertations and Theses (AAT 9527513).

McPherson, G. (1997). Giftedness and talent in music. *Journal of Aesthetic Education, 31(4),* 65-77.

Nease, A., Mudgett, B., & Quinones, M. (1999). Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. *Journal of Applied Psychology, 84(5),* 806-814.

O'Neill, S., & Sloboda, J. (1997). The effects of failure on children's ability to perform a musical test. *The Society of Research in the Psychology of Music and Music Education, 25,* 18-34.

Peterson, C., Semmel, A., Von Baeyer, C., Abramson, L., Metalsky, G., & Seligman, M. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research, 6(3),* 287-299.

Pintrich, P., & DeGroot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82(1),* 33-46.

Pulford, B., & Colman, A. (1997). Overconfidence: Feedback and item difficulty effects. *Personality and Individual Differences, 23(1),* 125-133.

Radocy, R. & Boyle, J. (2003). *Psychological foundations of musical behavior.* Springfield, Illinois: Charles C. Thomas, Publisher, Ltd.

Reeve, C., Bonaccio, S., & Charles, J. (2008). A policy-capturing study of the contextual antecedents of test anxiety. *Personality and Individual Differences*, 243-248.

Roberts, T. (1991). Gender and the influence of evaluations on self-assessments in achievement settings. *Psychological Bulletin, 109(2),* 297-308.

Schmidt, C. (1995). Attributions of success, grade level, and gender as factors in choral students' perceptions of teacher feedback. *Journal of Research in Music Education 43(4),* 313-329.

Schunk, D. (1989). Self-Efficacy and achievement behaviors. *Educational Psychology Review, 1*, 173-208.

Schunk, D. H. (1991). Self-Efficacy and Academic Motivation. Educational Psychologist, 26(3), 207-231.

Schunk, D. (2008). Attributions as motivators of self-regulated learning. In S. &. (ed.), *Motivation and self-regulated learning: Theory, research, and applications* (245-266.). New York: Lawrence Erlbaum Associates.

Seligman, M., Nolen-Hoeksema, S., Thornton, N., & Thornton, K. (1990). Explanatory style as a mechanism of disappointing athletic performance. *Psychological Science, 1(2),* 143-146.

Shute, V. (2008). Focus on Formative Feedback. *Review of Educational Research, 78,* 153-189.

Silver, W., Mitchell, T., & Gist, M. (1995). Responses to successful and unsuccessful performance: The moderating effect of self-efficacy on the relationship between performance and attributions. *Organizational Behavior and Human Decision Processes, 62(3),* 286-299.

Simons, G. (1976). *Manual for the Simons Measurement of Music Listening Skills (SMMLS) for Young Children.* Chicago: Stoelting Company.

Sloboda, J. (1985). *The Musical Mind: The Cognitive Psychology of Music.* Oxford: Clarendon Press.

Smith, S., Kass, S., Rotunda, R., & Schneider. (2006). If at first you don't succeed: Effects of failure on general and task-specific self-efficacy and performance. *North American Journal of Psychology, 8(1),* 171-182.

Steinmayr, R., & Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles, 61*, 736-749.

Tarrant, R., Leathem, J., & Flett, R. (2010). What have sport and music performance taught us about test anxiety? *Psychology Journal, 7(2),* 67-77.

Tzetsis, G., Votsis, E., & Kourtessis, T. (2008). The effect of different corrective feedback methods on the outcome and self-confidence of young athletes. *Journal of Sports Science and Medicine, 7*, 371-378.

Vargas-Tonsing, T. (2009). An exploratory examination of the effects of coaches' pre-game speeches in athletes' perceptions of self-efficacy and emotion. *Journal of Sport Behavior, 32(1),* 92-111.

Vrugt, A.; Langereis, M.; & Hoogstraten, J. (1997). Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *Journal of Experimental Education, 66(1),* 61-73.

Waldersee, R., & Luthans, F. (1994). The impact of positive and corrective feedback on customer service performance. *Journal of Organizational Behavior, 15(1),* 83-95.

Weber, C., & Bizer, G. (2006). The effects of immediate forewarning of test difficulty on test performance. *The Journal of General Psychology, 133(3),* 277-285.

Weinberg, R., Gould, D., Yukelson, D., & Jackson, A. (1981). The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance test. *Journal of Sport Psychology, 4*, 345-354.

Weiner, B. (1974). *Cognitive views of human motivation.* New York: Academic Press.

Wentzel, K. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development, 73(1),* 287-301.

Young, W. (1972). Musical Aptitude Profile norms for use with college and university nonmusic majors. *Journal of Research in Music Education*, 385-390.

Zimmerman, B. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82-91.

Appendix A

Script to be read to music majors and non-music majors

Mrs. Hutton is a doctoral student at OU studying music education. At your next class meeting, Mrs. Hutton will gather some data from this class. She is studying the effects of immediate retesting on music aptitude test performance. You will be given a consent form to sign if you choose to participate. There are no overt benefits to you for taking the test other than the fact that there will be two drawings: one for a $10 i-Tunes gift certificate, and one for a $25 Olive Garden gift card, and all participants will be offered candy as thanks for their participation. Testing will take most of the class period. While you are not required to participate, attendance will be taken that day and absences noted. I hope you will participate in the spirit of contributing to the field of music education.

Appendix B

University of Oklahoma

Institutional Review Board

Informed Consent to Participate in a Research Study

Project Title:	The Effect of Immediate Retesting on Musical Aptitude Performance
Principal Investigator:	Paula Hutton
Department:	Music Education

You are being asked to volunteer for this research study. This study is being conducted at Oklahoma Christian University. You were selected as a possible participant because the study involves university students enrolled in music classes.

Please read this form and ask any questions that you may have before agreeing to take part in this study.

Purpose of the Research Study

The purpose of this study is: to investigate whether taking a music aptitude test twice in quick succession will raise scores on the second taking of the test.

Number of Participants

About 250 people will take part in this study.

Procedures

If you agree to be in this study, you will be asked to do the following:

In this study you will be asked to do the following: You will take a 15-minute music aptitude test in which you will hear pairs of short melodies. The test is on an audio recording. You will determine if each set of melodies is exactly alike, or different tonally, or different rhythmically. While the tests are being rated, you will be asked to fill out a questionnaire providing demographic information. You

will learn your rating of “high,” “middle,” or “low” from the first test and will be given another attempt on the music test. When testing is over, there will be a drawing for a $10 i-Tunes gift card, a $25 Olive Garden gift card, and everyone will be offered candy as a thank-you for participating.

Length of Participation

The entire procedure will take place during one 50-minute class period.

This study has the following risks:

Any risks are minimal.

Benefits of being in the study are

None

Confidentiality

In published reports, there will be no information included that will make it possible to identify you. Research records will be stored securely and only approved researchers will have access to the records.

There are organizations that may inspect and/or copy your research records for quality assurance and data analysis. These organizations include the University of Oklahoma Music Education Department and the OU Institutional Review Board.

Compensation

You will not be reimbursed for your time and participation in this study. There will be a random drawing for a $10 i-Tunes gift card and a $25 Olive Garden gift card for each class and all participants will be offered candy as thanks.

Voluntary Nature of the Study

Participation in this study is voluntary. If you withdraw or decline participation, you will not be penalized or lose benefits or services unrelated to the study. If you decide to participate, you may decline to answer any question and may choose to withdraw at any time.

Contacts and Questions

If you have concerns or complaints about the research, the researcher(s) conducting this study can be contacted at

Paula Hutton (405) 425-5533; email: paula.hutton@oc.edu

Faculty advisor: Dr. Charlene Dell (405) 325-0168; email cdell@ou.edu

Contact the researcher(s) if you have questions or if you have experienced a research-related injury.

If you have any questions about your rights as a research participant, concerns, or complaints about the research and wish to talk to someone other than individuals on the research team or if you cannot reach the research team, you may contact the University of Oklahoma – Norman Campus Institutional Review Board (OU-NC IRB) at 405-325-8110 or irb@ou.edu.

You will be given a copy of this information to keep for your records. If you are not given a copy of this consent form, please request one.

Statement of Consent

I have read the above information. I have asked questions and have received satisfactory answers. I consent to participate in the study.

Signature	Date

Appendix C

Musical Aptitude Self-Efficacy Scale

The statements below describe different aspects of musical ability. In the column **Confidence**, rate how confident you are that you can do them **as of now**. Rate your degree of confidence by recording a number from 0 to 100 using the scale given below:

0	10	20	30	40	50	60	70	80	90	100
Cannot do at all				Moderately certain can do					High certain can do	

Confidence

(0-100)

I can determine subtle tonal differences in two similar melodies. _____

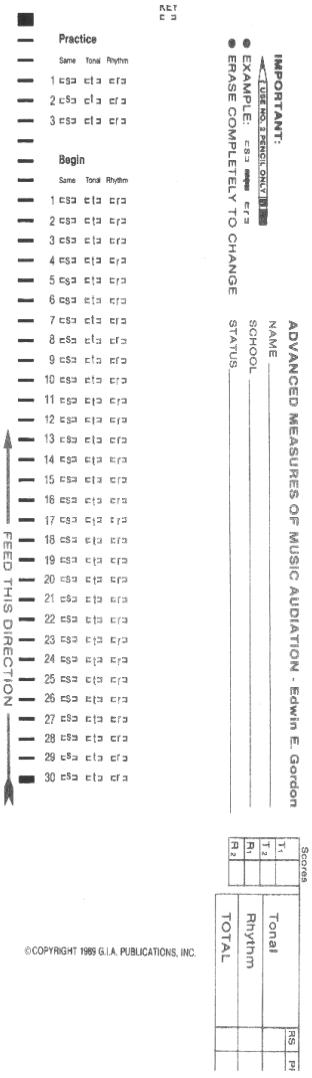
I can determine subtle rhythmic differences in two similar melodies. _____

I can determine that there are no differences between two melodies. _____

Please enter your code number here:_____

Appendix D

Sample Answer Sheet for Advanced Measures of Music Audiation



Appendix E

Questionnaire

Thank you for providing some demographic information. Should you change your mind about participating, you may withdraw from the study at any time. The information you provide will be coded and kept confidential.

Name_____

Your gender is:

_____ Male

_____Female

Your age is:

A. _____ younger than 18

B. _____ 18-23

C. _____ 24 or older

Your major:

A._____ Music Major

B. _____Not a Music Major

In the space below, please define musical ability in your own words. If musical ability is important to you, please describe how it impacts your life.

Appendix F

Random Number Sequence

Each even number corresponds to a positive rating (+).

Each odd number corresponds to a negative rating (-).

38	25	84	42	80
3	23	61	25	36
84	65	93	16	8
51	8	17	74	37
39	15	100	72	26
42	21	68	89	88
63	38	4	33	80
19	92	25	79	78
8	34	52	31	37
95	76	94	87	40
80	22	16	98	87
89	44	55	68	56
73	11	92	18	87
54	100	5	75	21
57	54	78	27	16
77	76	55	78	91
36	39	95	67	47
99	28	64	89	34
16	71	64	45	29
11	41	49	30	40

Appendix G

Manipulated Feedback Given to Students Prior to Posttest

(As seen on a Power Point Slide)

Please turn your test answer sheet over. On the back you will see a plus sign (+), a check mark (√), or a minus sign (--).

If you received a plus sign (+), you did extremely well on the pretest. Ratings in this category were at or near a perfect score.

If you received a check mark (√), your pretest score was in the middle range—neither very high nor very low.

If you received a minus sign (--), your pretest score was in the low- to very low range.

Appendix H

Debriefing Statement (OU)

Thank you for participating in my study. In psychological studies, it is sometimes necessary to conceal our hypotheses because when people know what is being studied they often alter their behavior. However, I do not want you to leave misinformed, so I will now tell you what I was actually studying.

The purpose of this study is to investigate the effect of positive and negative feedback on music aptitude test performance. I wanted to see if your rating changed the way you performed on the second test.

In order to test this hypothesis, I randomly assigned ratings on the first test. You read your false rating before you took the test a second time.

I apologize that I could not reveal my true hypothesis to you up front, but I hope you can see why it was necessary to keep this information from you. When people know exactly what the researcher is studying, they often change their behavior, thus making their responses unusable for drawing conclusions about human nature and experiences. **For this reason, I ask that you please not discuss this study with other students who might participate. Thank you for your cooperation.**

If your participation in this study has in any way upset you, please feel free to set up an appointment with one of OU's licensed psychologists or counselors. Counseling and Testing Services is located on the second floor of Goddard Health Center, and they can be reached at (405) 325-2911 or 325-2700.

If you have any questions about this study, feel free to ask Ms. Hutton at paula.hutton@oc.edu or phone 425-5533. Thank you for your help today.

Now that you know the true purpose of this study, please check this box *if you would like your data to be **excluded** from my study:*

_____ _____

Signature of Participant or Participant # Date

Appendix I

Institutional Review Board Approval Letter

The University of Oklahoma

OFFICE OF HUMAN RESEARCH PARTICIPANT PROTECTION - IRB

IRB Number: 13068
Approval Date: May 31, 2011

May 31, 2011

Paula Hutton
Dept of Music Education
1317 Harding Ave
Edmond, OK 73013

RE: The Effect of Immediate Retesting on Musical Aptitude Performance

Dear Ms. Hutton:

Thank you for completing and returning the IRB Application for Continuing Review (Progress Report) for the above-referenced study. You have indicated that the study is still active. I have reviewed and approved the Progress Report and determined that this study was appropriate for continuation.

This letter documents approval to conduct the research as described in:
Cont Review Form Dated: May 26, 2011
Protocol Dated: May 26, 2011
Consent form - Subject Dated: May 26, 2011

Please remember that any change in the protocol, consent document or other recruitment materials (adverstisements, etc.) must be approved by the IRB prior to its incorporation into the study procedures. Submit a completed Protocol Modification form to the IRB office.

Approximately two months prior to the expiration date of this approval, you will be contacted by the IRB staff about procedures necessary to maintain this approval in an active status. Although every attempt will be made to notify you when a study is due for review, it is the responsibility of the investigator to assure that their studies receive review prior to expiration.

The approval of this study expires on May 30, 2012 and must be reviewed by the convened IRB prior to this time if you wish to remain in an active status. Federal regulations do not allow for extensions to be given on the expiration date.

If we can be of further assistance, please call the IRB office at (405) 325-8110 or send an email to irb@ou.edu.

Cordially

Laurette Taylor, Ph.D.
Chair, Institutional Review Board

1616 West Lindsey, Suite 150 Norman, Oklahoma 73069 PHONE: (405) 325-8110