The Role of Calibration Bias and Performance Feedback in Achievement Goal Regulation

Krista R. Muis1*, Philip H. Winne2, John Ranellucci3

1 Department of Educational and Counselling Psychology, McGill University
2 Faculty of Education, Simon Fraser University
3 Department of Educational Psychology and Special Education, Michigan State University

* Correspondence concerning this article should be addressed to Krista R. Muis, McGill University, 3700 McTavish Street, Montreal, Quebec, H3A 1Y2, or via e-mail at krista.muis@mcgill.ca.

Received: October 25, 2015   Accepted: January 15, 2016   Online Published: April 6, 2016
DOI: 10.12735/ier.v4i1p14            URL: http://dx.doi.org/10.12735/ier.v4i1p14
Copyright © Krista R. Muis et al.**

Abstract
Do achievement goals change across time in response to performance feedback? Does goal orientation relate to calibration of estimated to actual achievement? We studied these issues over three tasks spanning a semester-long course where ninety-nine undergraduates received feedback about performance on each task. Learners were consistently and quite substantially biased in estimating performance with bias inversely related to actual performance. Goal orientation was not stable across time as a function of task, and it varied in some tasks in relation to calibration accuracy. These findings demonstrate goal orientations are sensitive to task and feedback. Moreover, goal orientation had varying and sometimes no relation to achievement, with calibration bias mediating most of the relations. In an authentic setting where learners experience multiple tasks over time, it is important to consider individuals’ calibration bias for performance on specific tasks. Calibration bias may be a key factor in learners’ regulation of achievement goals.

Keywords: achievement goal regulation, calibration, feedback, mastery goals, performance goals

1. Introduction
Over the past three decades, achievement goals have been shown to predict important educational outcomes (Ames, 1992; Hulleman, Schrager, Bodmann, & Harackiewicz, 2010). According to
achievement goals are defined as “a future-focused cognitive representation that guides behavior to a competence-related end state that the individual is committed to either approach of avoid” (p. 423). Current theories conceptualize achievement goals under a 2 x 2 (Elliot & McGregor, 2001; Elliot & Murayama, 2008) or trichotomous framework (Elliot & Church, 1997; Middleton & Midgley, 1997; Pintrich, 2000a). Under the 2 x 2 framework, adopted here, achievement goals focus on competence that is developed (mastery) or demonstrated (performance), and are further divided into approaching positive outcomes (approach) or avoiding negative ones (avoidance). This model distinguishes four kinds of achievement goals: mastery-approach, mastery-avoidance, performance-approach, and performance-avoidance (Elliot & McGregor, 2001).

According to Elliot and McGregor (2001), learners who adopt a mastery-approach orientation strive to develop competence and task mastery. They believe effort and outcome covary. Individuals who adopt mastery-approach goals focus on improving their skills or doing better than they have in the past. In this regard, intra-individual comparisons are made. In contrast to the mastery-approach orientation, learners who adopt a mastery-avoidance orientation focus on avoiding failure (relative to oneself, not in comparison to others). For the mastery-avoid goal construct, incompetence is the focus. A mastery-avoid oriented learner, for example, may strive to evade misunderstanding or failing to learn course material, strive not to forget what has been learned, or try not to lose previously developed physical or intellectual capabilities (Elliot & McGregor, 2001). An exemplar is a perfectionist who tries not to make any mistakes (Pintrich, 2000b). Conceptually, the mastery component in mastery-avoid goal orientation emerges from optimal antecedents (e.g., motive dispositions, implicit theories, socialization histories) that may facilitate positive consequences (e.g., mastery-approach goals; see Elliot & McGregor, 2001, for a complete discussion). The avoidance component, however, is hypothesized to emerge from non-optimal antecedents and may result in negative consequences (e.g., poor performance or high levels of anxiety). Several psychometric studies support the 2 x 2 dimensional framework of achievement goal orientation (e.g., Elliot & McGregor, 2001; Finney, Pieper, & Barron, 2004; Muis & Winne, 2012).

The third achievement goal is a performance-approach goal orientation. It characterizes learners who strive to demonstrate aptitude and seek favorable judgments. Individuals who adopt a performance-approach orientation may try to look good compared to others (the appearance component) and/or to outperform others (the normative component) (Urdan & Mestas, 2006). A performance-approach goal orientation theoretically is associated with a belief that ability is necessary for success, effort does not improve performance and is interpreted as evidence of low ability—if an individual is able, then effort is unnecessary for success. In contrast, a performance-avoid orientation is rooted in a fear of failure, worry, and concern (Elliot & Church, 1997; Hulleman et al., 2010). Individuals who adopt a performance-avoid orientation focus on avoiding negative outcomes such as performing poorly or doing poorly compared to others (Hulleman et al., 2010).

Most research has focused on how the different types of achievement goals relate to various learning and motivational processes and outcomes, and how different classroom climates interact with or influence learners’ achievement goals. Reviews of the achievement goal literature suggest that the two avoidance goals are correlated to negative outcomes such as anxiety, disinterest, and low levels of achievement. In contrast, mastery-approach goals are generally positively correlated to adaptive motivational and cognitive processes and outcomes such as high interest and effective self-regulatory processes, whereas patterns of relations between performance-approach goals and various processes and outcomes are mixed (Ames, 1992; Baranik, Stanley, Bynum, & Lance, 2010; Elliot, 2005; Senko, Hulleman, & Harackiewicz, 2011).

1.1. Goal Orientation and Feedback

Although initial theories described how achievement goals might remain stable over tasks or contexts (e.g., endorsing the same goal or level of a particular goal over time) or change within an
academic context (e.g., endorsing a different goal or changing the intensity of the level of the same goal), few studies have explored this aspect of achievement goal theory. Those that evaluated this assumption adopted a self-regulatory perspective to describe goal regulation (e.g., Fryer & Elliot, 2007; Senko & Harackiewicz, 2005; Muis & Edwards, 2009). Self-regulated learning refers to how individuals monitor and control their cognitive, metacognitive, and motivational processes within educational settings (Muis, 2007; Pintrich, 2000b; Winne & Hadwin, 1998; Zimmerman, 2000). Typically, the focus of goal regulation has been in terms of Locke and Latham’s (1990) definition of a goal, but recent theories of self-regulated learning have proposed that achievement goals may also be regulated during task engagement and/or over the course of several tasks (Muis, 2007; Pintrich, 2000b; Winne & Hadwin, 2008).

How do learners regulate their goals? Senko and Harackiewicz (2005) suggested two ways. First, learners may engage in goal switching, where they change from a mastery-approach goal to a mastery-avoidance (or vice versa), or from a mastery-approach goal to a performance-approach goal (or vice versa), or some combination thereof. This may occur over various tasks wherein an individual adopts a mastery-approach goal for a class paper but then switches to a performance-approach goal for a midterm exam. The second type of goal regulation is goal intensification. It occurs when learners increase or reduce the level of goal endorsement without switching the type of goal being pursued. For example, a learner may highly endorse a mastery-approach goal for a class paper but only moderately endorse that same goal for a midterm exam. The same goal is being pursued but to a lesser extent. The change in the level of goal pursuit may be small (relatively stable) or large (unstable).

To date, studies that explored the stability of achievement goals focused on how goals: vary as a function of changes in task (e.g., Fryer & Elliot, 2007; Muis & Edwards, 2009); shift during transitions from elementary to middle school (Anderman & Anderman, 1999); or change as a function of feedback (Senko & Harackiewicz, 2005). For example, Senko and Harackiewicz (2005) examined whether learners engaged in goal switching or goal intensification after receiving performance feedback on a series of tasks. In Study 1, conducted in a college classroom, students’ goal pursuits remained stable throughout the semester but poor exam performance predicted a significant decrease in mastery-approach and performance-approach goal pursuit. In Study 2, a laboratory-based study, individuals were given “false” feedback that indicated they were “well above average” or “well below average” on a set of tasks. Results revealed negative feedback reduced individuals’ level of mastery-approach goals and performance-approach goals were related to higher levels of achievement on the task.

Why might these patterns arise? First, feedback provides information about competence or incompetence, and goal orientation influences how that information is interpreted (Bobko & Coella, 1994). For example, in the face of negative feedback, individuals with a mastery-approach goal orientation theoretically view feedback as useful in guiding future behaviour to increase mastery (and performance). They may increase effort or revise tactics and strategies (Dweck & Leggett, 1988; Elliott & Dweck, 1988). Individuals with a performance-approach or avoid goal orientation theoretically would be more likely to experience negative affect and attribute failure to lack of ability (Dweck & Leggett, 1988). Elliot and Harackiewicz (1996) proposed that individuals with high performance avoiding orientation focus on failure-relevant information that, in turn, may depress subsequent performance as fear of failure increases. Moreover, because performance-oriented individuals emphasize managing impressions over developing competency, positive feedback does not spur them to pursue more challenging goals because this path increases the risk of future negative feedback (VandeWalle, Brown, Cron, & Slocum Jr., 1999).
1.2. Goal Orientation and Calibration

We argue that the type of feedback individuals receive has an important antecedent not yet explored in the literature. That is, how individuals react to performance feedback depends on their initial expectations with regard to their performance on a specific task; specifically, they may be overconfident or underconfident in their predictions. From a self-regulated learning perspective, this is an important component that may predict goal regulation. Moreover, individuals’ goal orientations may, in turn, predict their calibration. Calibration is the degree (accuracy) and direction (bias) of discrepancy between a student’s perception about learning and actual properties of learning (Jamieson-Noel & Winne, 2003; Winne & Jamieson-Noel, 2002). Although self-enhancement and self-evaluation motives theoretically influence calibration, Brown (1990) suggested individuals’ implicit beliefs about ability must be considered if one is to understand which influence is predominant. According to Brown and consistent with social comparison theory (Festinger, 1954), even though individuals favor positive reinforcement (e.g., which suggests high ability), they do not completely refrain from seeking feedback that may disclose incompetence. Accordingly, two motivational influences operate in calibration: self-evaluation, to obtain accurate evaluation information; and self-enhancement, to increase one’s perception of competence. Individuals must resolve tension between acquiring information for its instrumental value and wanting to protect one’s ego and self-esteem (Northcraft & Ashford, 1990).

In this context, Wahlstrom (2001) proposed a parallel between a performance orientation and self-enhancement, and a mastery orientation and self-evaluation. Since mastery-oriented learners are concerned with improving competence, they seek feedback to develop competence and are likely to remain unbiased in estimating performance even when expectations are low. Since the value of feedback outweighs potential costs (e.g., lowered status), mastery-oriented learners need not modify perceptions of their abilities. As well, mastery-oriented learners report greater use of learning strategies and reports of self-regulation (Middleton & Midgley, 1997; Wolters, Yu, & Pintrich, 1996), which some theorists believe lead to better calibration accuracy. In contrast, performance-approach oriented learners strive to display competence and outperform others. Thus, they must set relatively high goals, which expresses overconfidence. If they perceive success is likely, they are even more likely to overestimate performance. Martin and Debus’s (1998) research supported this hypothesis. Students with an ego-orientation held high mathematics self-concepts and overrated their perceptions of ability compared to others.

Performance-avoid oriented learners strive to avoid displaying lack of ability. Feedback for these individuals can come at a great cost, so they are likely to underestimate performance to avoid negative feedback that demonstrates lack of ability (Wahlstrom, 2001). This logic is consistent with Elliot and Church’s (1997) proposal whereby performance-avoid learners are rooted in a fear of failure and hold low expectations of their competence. Phillips’ (1984) research supports this. Children who believed they had low ability displayed lower achievement standards than children with average or higher beliefs about competence. Children who believed their competence was low mainly underestimated their actual performance. Phillips argued these children underestimated performance to protect judgments of competence from internalized sources of criticism and feedback.

To test these hypotheses, Wahlstrom (2001) studied secondary students who received feedback on their skills in a manufacturing technology program to examine whether goal orientation predicted students’ perceptions of the assessment process—fairness, utility, and perceived threat—and their post-assessment motivation: self-efficacy and intentions to seek feedback. He also examined whether goal orientation predicted the accuracy of students’ calibration (computed as actual score minus predicted score). He found that high mastery-approach goal orientation correlated positively with high self-efficacy and intentions to seek feedback. Holding a
performance-approach goal orientation also correlated positively with high self-efficacy, but a performance-avoid goal orientation did not predict students’ post-assessment motivation. Calibration bias and judgments of utility mediated these relationships depending on the skill investigated. Performance-avoid goals did not correlate with calibration but both a mastery-approach orientation and a performance-approach orientation negatively correlated with calibration. That is, students with higher levels of either mastery- or performance-approach goal orientation were less likely to underestimate their performance.

Based on Wahlstrom’s (2001) model and results from Martin and Debus (1998) and Phillips (1984), we make the following predictions. First, we predict mastery-approach individuals will accurately predict performance across several tasks. Thus, we expect no relationship between a mastery-approach goal orientation and calibration. Second, we predict that a mastery-avoid goal orientation will negatively relate to calibration; the higher learners’ mastery-avoid goal orientation, the more they will underestimate performance across all tasks. We also predict a performance-approach goal orientation will be positively related to calibration; the higher learners’ performance-approach goal orientation, the greater the over-estimation in performance. We expect this relationship to hold across three tasks in our study. Moreover, we predict performance-avoid goal orientation will negatively related to calibration such that performance-avoid individuals will be more likely to underestimate their performance across all three tasks. Given that we predict a similar relationship between learners’ goal orientation and calibration across the three tasks, we also predict learners’ calibration across tasks will be positively related. Finally, we predict calibration bias will be positively related to performance on each task.

1.3. The Present Study

We believe research relating goal orientation and performance feedback has several shortcomings. Specifically, with the exception of Muis and Edwards (2009), most studies measured goal orientation at just one or two times, typically, the beginning and end of the study or course. Second, with the exception of Wahlstrom (2001), studies have not taken into consideration how calibration may mediate relations between initial goal orientations, actual performance, and subsequent goal orientations. Given that calibration is a significant factor in models of self-regulated learning (Winne, 2005), calibration should also be considered in achievement goal regulation as how one reacts to feedback depends on their initial performance expectations coupled with their goal orientations. Moreover, if goals change as feedback varies over time across tasks, subsequent goal orientation may be more predictive of following performance than initial goal orientation. Accordingly we investigated goal orientation for achievement in an undergraduate course as a function of external evaluations students received on successive course assignments. Our study addresses three related research questions: (1) Are achievement goals stable across time or do they vary as function of response to feedback? (2) What is the relationship between achievement goal orientation and calibration bias? (3) Does calibration bias mediate relations between achievement goals and performance? We obtained data on goal orientation at four times over the semester in an undergraduate course using Elliot and McGregor’s (2001) Achievement Goal Questionnaire (AGQ) as a representation of their 2 x 2 model of goal orientation—mastery versus performance orientation by approach versus avoidance focus.

Since we predict mastery-approach individuals to be relatively accurate in predicting performance, we do not expect to see changes in mastery-approach goal orientation. Even in the face of negative feedback, mastery-approach individuals will use that information to improve competence and will not adjust goals. Thus, a mastery-approach goal orientation should remain stable across tasks after feedback. In contrast, since mastery-avoid individuals fear failing and are likely to underestimate performance, we predict that if mastery-avoid individuals receive positive
feedback, they will become less avoidant. Thus, we expect to see a decrease in mastery-avoidance across tasks after feedback. If, however, they receive negative feedback (though we suspect this will not occur since they are more likely to underestimate rather than overestimate performance), mastery-avoid individuals will become more avoidant, and an increase in mastery avoidance across tasks will result.

Because performance-approach individuals are more likely to overestimate performance, they are more likely to receive negative feedback. With negative feedback, performance-approach individuals may be threatened by future failure and, as such, may lower goals. Thus, we predict that a performance-approach goal orientation will decrease over feedback episodes. Finally, for performance-avoid individuals, since they are more likely to underestimate performance to avoid negative feedback, they are more likely to receive positive feedback. This positive feedback will result in performance-avoid individuals lowering their avoidance goals. Thus, we predict a decrease in performance-avoid goal orientation over feedback episodes.

In sum, we expect mastery-approach goals to remain stable and all other goal orientations to decrease over multiple feedback episodes. Moreover, although we expect changes in levels of goal orientation as a function of feedback, we do not expect learners to completely abandon a particular goal type. That is, we expect goal intensification (in the form of a reduction) but not goal switching. Thus, we predict that prior goal orientation will be positively related to subsequent goal orientation across all four measures. Moreover, based on Hulleman et al.’s (2010) meta-analysis of the literature, we predict that mastery-approach goals will not predict performance but that performance-approach goals will positively predict performance, which will be positively mediated by calibration bias. Both mastery-avoidance and performance-avoidance goals will be negatively related to performance, and calibration bias will negatively mediate relations between these goals and performance. Finally, feedback (as a function of performance) will positively predict subsequent performance-approach goals but negatively predict mastery-avoidance and performance-avoidance goals.

2. Method
2.1. Participants
Participants were 99 students, 13 males and 86 females, taking an educational psychology course. This represents 40% of total course enrollment and a ratio of males to females common in education courses at that university. All participants received and signed a university research ethics board approved consent form. The mean age of students was 23.12 years (SD = 6.97), and the mean self-reported GPA was 3.42 (SD = 2.96, N = 85).

2.2. Measures
2.2.1. Goal Orientation
We used the 12-item Achievement Goal Questionnaire (AGQ; Elliot & McGregor, 2001) to assess students’ achievement goals for their educational psychology course. Students indicated their agreement with each statement on a scale of 1 (not at all true of me) to 7 (very true of me). The AGQ generates four a priori non-overlapping subscales of three items each: mastery-approach, mastery-avoidance, performance-approach and performance-avoidance. Responses to the three items within each subscale were averaged. Sample items are: “I desire to completely master the material presented in this class” (mastery approach), “I worry that I may not learn all that I possibly could in this class” (mastery avoid), “It is important for me to do better than other students”
(performance approach), and “My goal in this class is to avoid performing poorly” (performance avoid). Reliability estimates from our sample across the four times ranged from .67 to .96.

2.2.2. Calibration Bias and Accuracy
Calibration bias was measured by subtracting students’ actual score on each task from their predicted score. Calibration accuracy is the absolute value of the bias score. Bias indicates the magnitude of judgment error between students’ expectation for performance and actual performance. It reflects overconfidence (a positive score) or underconfidence (a negative score). Because the goal orientation groups were expected to over- or under-estimate performance, we used bias as a variable in our models rather than accuracy.

We acknowledge that calibration bias is mathematically dependent on performance and, thus, the two measures are inherently correlated. Statistically, all measures of calibration discussed in the self-regulated learning literature share this property. Nonetheless, calibration is theoretically critical in models of self-regulated learning since it is the signed (directional) discrepancy between expectations (goals) and performance that triggers self-regulating learning (Winne, 1995; Winne & Perry, 2000). Given that only one study has examined relations between calibration bias and performance (e.g., Wahlstrom, 2001), we are still uncertain whether calibration bias is negatively or positively related to performance and whether that relationship is task specific. Our models examine this issue under the constraint that performance and calibration bias are inherently correlated.

2.3. Procedure
In this course, students wrote two short papers (5 pages) that required them to reflect on course content and describe how it relates to teaching. They also took multiple-choice midterm and final exams. In week 2 of the semester, before participants began work on papers or intense studying for a test, they estimated their course grade by checking one of eleven percent ranges (e.g., 92-100%) listed next to a corresponding letter grade ranging from A+ to F. Next, they estimated performance on each of paper 1, think paper 2, midterm exam, and final exam in three ways: “What percent will you try to achieve?” “What do you think is the lowest possible percent you might actually receive?” “What do you think is the highest possible percent you might actually receive?” Participants then completed the AGQ.

When students handed in each paper and each exam, they estimated performance again using the scoring scale for that assignment: 20 points for the papers, 30 points for the midterm multiple-choice exam, and 60 points for the final multiple-choice exam. They also reported the score they tried to achieve (their goal) and the lowest and highest possible scores they believed they might actually receive.

When papers and the midterm exam were handed back, students responded to another questionnaire. To insure they attended to the grade they received and the psychological context of receiving that grade, they first reported the score they had previously estimated they would receive when they handed in the assignment and then reported the score they actually received on the assignment. To make salient the nature of feedback provided by markers of the papers, we asked students to rate the marker’s feedback (very helpful, helpful, somewhat helpful, not helpful at all, or didn’t look at feedback) and to “explain how feedback … helps you in the course. If you think the feedback won’t help, explain why.” There was no written feedback about the midterm exam, so these questions were not appropriate. Finally, on each of these three occasions, students again completed the AGQ after the foregoing activities. Instructions for responding to the AGQ were the same across all tasks. Students were asked to indicate how well each statement best describes them. No specific reference was made to each task to enhance students’ focus on the feedback they received.
2.4. Preliminary Report and Overview of Statistical Analyses

To examine the first research question (the stability of goal orientation), we computed a $4 \times 4$ doubly multivariate repeated measures ANOVA. Based on our hypotheses, we expected a main effect of time, which would suggest that goal orientations change across contexts and as a function of feedback. We also examined stability across contexts as a function of relations between each respective goal orientation across two time points in a path model. Low correlations across time points would suggest goal orientations are unstable. Based on theoretical considerations and previous research (Dweck & Leggett, 1988; Fryer & Elliot, 2007; Muis & Edwards, 2009; Senko & Harackiewicz, 2005), however, we expected some degree of stability in individuals’ achievement goals. For the second and third research questions, we employed path modeling using EQS (Bentler & Wu, 1995) to model relations between achievement goals, calibration bias, and achievement to examine whether relations varied as a function of context and feedback about achievement.

3. Results

Data obtained in week 2 of the semester is henceforth labeled “time 1” or from the “start of the course.” We first examined subscales of the AGQ for normality. Kline (1998) suggested using absolute cut-off values of 3.0 for skewness and 8.0 for kurtosis. All items on the AGQ were well within these ranges (ranging from –1.79 to 0 for skewness and from –1.18 to 4.88 for kurtosis).

3.1. Changes in Goal Orientation as a Function of Feedback

To investigate changes in goals after receiving feedback (Research Question 1), we computed a $4 \times 4$ doubly multivariate repeated measures ANOVA. We present in Table 1 descriptive statistics and alpha internal consistency reliability coefficients for each goal orientation subscale at four time points: start of the course, and after receiving and explicitly attending to their expected score and actual score (feedback, the grade received) on each of the first paper, the midterm, and the second paper. Table 2 displays the descriptive statistics and calibration bias and accuracy scores for each of the course tasks.

Table 3 presents correlations between all variables. Finally, means for each goal orientation subscale across time are also shown in Figure 1.

A statistically detectable main effect was found for goal orientations, $F(3, 41) = 8.03$, $p < .01$, $\eta^2 = .16$; and for time, $F(3, 41) = 3.71$, $p < .01$, $\eta^2 = .08$. The interaction between time and orientations was not statistically different from zero $F(9, 41) = .97$, $p > .10$. Thus, over all four points in time, students had a higher mastery approach orientation, followed by mastery avoid, performance avoid, and performance approach. The main effect of time indicates that students’ overall orientation to goals changed in response to feedback and the task that feedback addressed. Specifically, as depicted in Figure 1 and as we predicted, students’ goal orientations decreased over time between the first and last measures of goal orientation. The absence of a statistically detectable interaction indicates that goal orientations changed over time only in degree (goal reduction) and not in kind (goal switching).

Post hoc comparisons using the LSD procedure were calculated to identify differences over time in overall goal orientation, and to assess differences in specific goal orientations collapsed across times. Overall goal level did not differ between times one through three (all $p > .05$) but that overall goal levels at each of times 1, 2, and 3 differed from time 4 (all $p < .001$). Moreover, overall mastery-approach orientation was statistically stronger than performance-approach and performance-avoid orientations (both $p < .05$) but not different from mastery-avoid orientation ($p > .05$). Finally, performance-approach orientation was statistically smaller than mastery-avoid ($p < .05$). No other statistically detectable differences were found.
### Table 1. Descriptive Statistics and Cronbach αs for Goal Orientations

<table>
<thead>
<tr>
<th>Scale</th>
<th>Time 1 Start of Course</th>
<th>Time 2 Think Paper 1</th>
<th>Time 3 Midterm</th>
<th>Time 4 Think Paper 2</th>
<th>Overall Orientation Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>α</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Mastery Approach</td>
<td>5.77</td>
<td>.89</td>
<td>.70</td>
<td>5.54</td>
<td>1.12</td>
</tr>
<tr>
<td>Mastery Avoid</td>
<td>4.26</td>
<td>1.49</td>
<td>.85</td>
<td>4.34</td>
<td>1.44</td>
</tr>
<tr>
<td>Performance Approach</td>
<td>4.36</td>
<td>1.40</td>
<td>.90</td>
<td>4.31</td>
<td>1.57</td>
</tr>
<tr>
<td>Performance Avoid</td>
<td>4.47</td>
<td>1.43</td>
<td>.67</td>
<td>4.36</td>
<td>1.49</td>
</tr>
<tr>
<td>Overall Goal Level</td>
<td>4.72</td>
<td></td>
<td></td>
<td>4.64</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. Descriptive Statistics and Calibration Bias and Accuracy on Think Papers 1 and 2, and Midterm Exam

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Think Paper 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance (as % of maximum)</td>
<td>82.98</td>
<td>8.82</td>
</tr>
<tr>
<td>Bias (as % of scale)</td>
<td>17</td>
<td>43</td>
</tr>
<tr>
<td>Accuracy (as % of scale)</td>
<td>37</td>
<td>27</td>
</tr>
<tr>
<td><strong>Think Paper 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance (as % of maximum)</td>
<td>82.33</td>
<td>8.91</td>
</tr>
<tr>
<td>Bias (as % of scale)</td>
<td>19</td>
<td>56</td>
</tr>
<tr>
<td>Accuracy (as % of scale)</td>
<td>43</td>
<td>41</td>
</tr>
<tr>
<td><strong>Midterm</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance (as % of maximum)</td>
<td>76.72</td>
<td>14.37</td>
</tr>
<tr>
<td>Bias (as % of scale)</td>
<td>28</td>
<td>45</td>
</tr>
<tr>
<td>Accuracy (as % of scale)</td>
<td>38</td>
<td>37</td>
</tr>
</tbody>
</table>

**Note:** We were not able to obtain final exam scores, so we could not calculate bias and accuracy for that task.
Table 3. Correlations between Each Goal Orientation, Calibration Bias, and Performance across Tasks

<table>
<thead>
<tr>
<th></th>
<th>Think Paper 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>MAP_1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAV_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-.43**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAP_3</td>
<td>-.05</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-.01</td>
<td>.41**</td>
<td>.32**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias_5</td>
<td>-.11</td>
<td>-.02</td>
<td>-.01</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>Performance_6</td>
<td>.23*</td>
<td>-.04</td>
<td>.07</td>
<td>-.17</td>
<td>-.70**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Midterm</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>MAP_1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAV_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.56**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAP_3</td>
<td>.03</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-.04</td>
<td>.39*</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias_5</td>
<td>.16</td>
<td>.28*</td>
<td>-.24</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>Performance_6</td>
<td>-.04</td>
<td>-.24*</td>
<td>.32*</td>
<td>-.34**</td>
<td>-.86**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Think Paper 2</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>MAP_1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAV_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.50**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAP_3</td>
<td>-.07</td>
<td>-.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-.10</td>
<td>.39**</td>
<td>.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias_5</td>
<td>-.25</td>
<td>-.17</td>
<td>.06</td>
<td>-.15</td>
<td></td>
</tr>
<tr>
<td>Performance_6</td>
<td>.25</td>
<td>.20</td>
<td>.01</td>
<td>.01</td>
<td>-.76**</td>
</tr>
</tbody>
</table>

Note: MAP = mastery approach, MAV = mastery avoidance, PAP = performance approach, PAV = performance avoidance, Bias = calibration bias, * p < .05, and ** p < .01.
3.2. Relations between Goal Orientation and Calibration of Achievement

Because the first set of analyses cannot disentangle whether goals changed as a function of task or feedback, we used EQS (Bentler & Wu, 1995) to examine the role of feedback in learners’ regulation of achievement goals. We also calculated parameter estimates in a series of path models to test our hypotheses about relations among goal orientation, calibration, and performance. Individual models were assessed at each time to increase power. (A latent growth curve model would have been preferable but we were not able to examine that model due to sample size.) For each model, we included goal orientation measured before a focal task, that is, from the preceding time, and after feedback was received on the focal task. Thus, two temporally adjacent measurements of goal orientation were included in each model. Specific calibration bias for the relevant task and bias for the following task were also included. The hypothesized model is displayed in Figure 2.

We tested for mediation according to Preacher and Hayes’ (2008) bootstrapping technique (MEDIATE) which is recommended when sample sizes are smaller as this method maintains higher levels of power while controlling for Type I errors. Specifically, calibration bias for each task was investigated as a mediator between each of the achievement goals and performance on that task. We report results from each test in the respective sections below. The three final models are presented in Figures 3, 4, and 5. For clarity, the figures include only statistically detectable paths (p ≤ .05) gauged using standardized coefficients.
Figure 2. General Structural Equation Model

Note: MAp = master approach, MAv = mastery avoid, PAp = performance approach, PAv = performance avoid, Bias A, B = calibration bias for specific task, score = performance on task A. + = predicted positive relationship, - = predicted negative relationship, and φ = no predicted relationship
Figure 3. Structural Equation Model for the First Think Paper

Note: MAP = master approach, MAV = mastery avoid, PAP = performance approach, PAV = performance avoid, bias = calibration bias for think paper (1) and midterm (2), Score = performance on first think paper.
Figure 4. Structural Equation Model for Midterm

Note: MAp = master approach, MAv = mastery avoid, PAp = performance approach, PAv = performance avoid, Bias = calibration bias for midterm (2) and second think paper (3), Score = performance on midterm.
Figure 5. Structural Equation Model for Second Think Paper

Note: MAp = master approach, MAV = mastery avoid, PAp = performance approach, PAv = performance avoid, Bias 3 = calibration bias for second think paper, Score = performance on second think paper.
### 3.2.1. First Think Paper

For the first think paper, estimation of the hypothesized model resulted in a good fit, $\chi^2 (40) = 75.46$, $p < .01$, CFI = .93, and RMSEA = .08. Bootstrap results revealed that mediation was not detectable for any of the hypothesized paths. As such, all paths in Figure 3 are direct effects. As predicted, initial mastery-approach goal orientation was unrelated to calibration for achievement (.00, $p > .05$). Unexpectedly, mastery-approach goal orientation was a positive predictor of performance (.28, $p < .05$). Also as predicted, mastery-avoidance negatively predicted calibration bias (-.16, $p < .05$). Contrary to our predictions, performance-approach negatively predicted calibration bias (-.15, $p < .05$), whereas performance-avoidance positively predicted calibration bias (.22, $p < .05$). Moreover, each of the initial goal orientation dimensions related strongly and proportionally to subsequent goal orientations: .72 for mastery-approach, .76 for mastery-avoid, .79 for performance-approach, and .63 for performance-avoid (all $p < .01$). Finally, calibration bias for the first think paper was positively related to calibration bias for the midterm exam (.23, $p < .01$).

As predicted, there was a statistically detectable relationship between calibration bias and actual performance on the first think paper. The direction of the relationship was, however, opposite to our prediction. Specifically, the more learners overestimated performance on the think paper, the lower they scored, whereas the more learners underestimated performance, the higher they scored (-.68, $p < .01$). These results are consistent with Wahlstrom’s (2001) study in which participants who overestimated performance did worse while those who underestimated did better. After feedback on this first think paper, there was a negative relationship between performance and subsequent performance-avoidance (-.27, $p < .01$). The lower individuals scored on the think paper, the higher their subsequent performance avoidance. No other relationships were statistically detected between feedback and subsequent goal orientation. Finally, three of the subsequent goal orientation dimensions were related to calibration bias on the midterm exam. Contrary to our predictions, the stronger learners’ mastery-avoid goal orientation the more they overestimated performance on the midterm exam (.31, $p < .01$). Similarly, the stronger learners’ performance-avoidance the more they overestimated performance on the midterm (.18, $p < .05$). In contrast, the stronger learners’ performance-approach goal orientation the more they underestimated performance on the midterm exam (-.28, $p < .01$).

### 3.2.2. Midterm Exam

For the midterm exam, estimation of the hypothesized model resulted in a good fit, $\chi^2 (37) = 99.16$, $p < .01$, CFI = .92, and RMSEA = .07. Unlike the first think paper, mastery-approach was not related to performance but, consistent with our predictions and results from the first think paper, it was also not related to calibration bias. Moreover, as was the case for the first think paper, mastery-avoidance was positively related to calibration bias (.31, $p < .05$) as was performance-avoidance (.18, $p < .05$). Performance-approach orientation, in contrast, was negatively related to calibration bias (-.30, $p < .05$). Bootstrap results statistically detected mediation for both performance goals. The total indirect effect of performance-approach orientation on performance through calibration bias was statistically detectable, with a point estimate of 1.81 and a 95% confidence interval of 1.3 to 4.29. The indirect effect of performance-avoidance on performance through calibration bias also was detected with a point estimate of -1.38 and a 95% confidence interval of -3.52 to -.26. No statistically detectable mediation effect was found for mastery-avoidance.

Consistent with results from the first think paper, calibration bias was negatively related to performance on the midterm (-.82, $p < .01$). The more learners’ overestimated performance (i.e., were overconfident) the lower they scored on the exam. In contrast, the more they underestimated performance the better they scored on the midterm. Similar to the patterns found for the first think paper, all goal orientation dimensions were positively related to subsequent goal orientation: .85 for mastery approach, .87 for mastery avoidance, .90 for performance approach, and .87 for
performance avoidance (all ps < .01). The relationship between feedback and subsequent performance-avoid goal orientation diminished to a non-statistically detectable level. There was, however, a positive relationship between feedback in the form of the midterm exam score and performance-approach orientation (.13, p < .05). The higher learners scored on the midterm exam the higher they rated their performance-approach goal orientation. Moreover, three of the subsequent goal orientations were related to calibration bias on the second think paper (the third task). Contrary to our predictions, mastery-approach orientation was negatively related to calibration bias on the second think paper (-.28, p < .05), as was performance-avoidance orientation (-.22, p < .05). In contrast, performance-approach orientation was positively related to calibration bias on the second think paper (.14, p < .05). Finally, no statistically detectable relationship was found between calibration bias for the midterm and calibration bias for the second think paper. This suggests that calibration bias may be task specific; learners were not consistently biased in their predictions of performance.

3.2.3. Second Think Paper

For the final task, the second think paper, estimation of the hypothesized model resulted in a good fit, χ² (32) = 79.72, p < .01, CFI = .94, and RMSEA = .08. Similar to results found for the midterm, mastery approach orientation was not related to performance. It was, however, negatively related to calibration bias for the second think paper (-.36, p < .01). The stronger learners’ mastery-approach orientation, the more they underestimated performance on the second think paper. Performance-avoidance orientation was also negatively related to calibration bias (-.25, p < .05), whereas performance-approach orientation was positively related (.14, p < .05). Bootstrap results revealed no statistically detectable mediation through calibration bias between each of the goals and performance on the second think paper.

Also consistent with the first model, calibration bias was negatively related to performance (-.79, p < .01). Also, goal orientation measured prior to performance was positively related to goal orientation after feedback on the second think paper, .83 for mastery-approach, .79 for mastery-avoidance, .83 for performance-approach, and .88 for performance-avoidance (all ps < .01). Finally, feedback (performance) on the second think paper positively predicted learners’ subsequent performance-approach orientation (.16, p < .05). No other statistically detectable relations were found. (Bias for the final exam was not obtained as we did not collect final exam scores.)

4. General Discussion

We examined relations among self-reports of goal orientation and calibration of achievement measured by a comparing a self-reported estimate of performance to actual performance. We also examined whether relations among these variables varied across types of tasks that were assigned across time, and whether feedback about performance on those tasks altered the magnitude of students’ expressed levels of goal orientation.

4.1. Calibration and Performance

Our study unambiguously demonstrates that learners are consistently biased in calibrating their performance (see also Winne & Jamieson-Noel, 2002). Specifically, learners who underestimate performance score higher on a task whereas learners who overestimate performance score lower. The average magnitude of bias is considerable, approximately 20% of the length of the scale used to grade the task (Table 3). As well, to the extent that calibration bias can be considered an influence on whether learners change studying based on metacognitive assessments of learning and acknowledging our non-experimental design, this bias has a substantial influence on performance. The standardized coefficients relating bias to performance in Figures 3, 4, and 5 are large both
statistically and practically. We note again, however, while calibration bias and performance are inherently dependent, excessive multicollinearity did not characterize our data. Thus, we submit that the negative relationship between the two constructs is important because it theoretically influences choices students make about study tactics they use to learn and, thus, what and how well they learn. We suggest that relations between calibration and performance need further exploration. We urge future research to investigate the effects of miscalibration bias on studying and whether reducing overconfidence and improving accuracy can benefit achievement and motivation.

4.2. Calibration and Goal Orientation

Consistent with our predictions, mastery-approach orientation was not related to calibration bias for the first two tasks (first think paper and midterm exam). However, individuals high on mastery-approach orientation underestimated performance on the second think paper. Results were also mixed for mastery-avoidance. Individuals high on mastery-avoidance underestimated performance on the first task but overestimated performance on the second task. Similarly, individuals high on performance-approach underestimated performance on the first and second tasks but overestimated performance on the third task. In contrast, individuals high on performance-avoidance overestimated performance on the first two tasks but then underestimated performance for the third task.

To interpret these results, for mastery-approach, we speculate that mastery-approach oriented learners strive to learn as much as possible but, consistent with this goal orientation, they perceive there is always more to learn and competence can be improved. It also could be the case that only through repeated feedback about their miscalibration did mastery-approach oriented learners realize they appropriateness of adjusting their performance expectations. In other words, mastery-approach learners may underestimate performance believing there is “room for improvement.” Additionally, as Dweck and Elliott (1983) proposed, mastery-approach individuals do not necessarily set high performance standards. Rather, the standards they set for learning include “focusing on learning” from a self-referential standard, which may not be too difficult to achieve (Hulleman et al., 2010).

We hypothesized that individuals high on mastery-avoidance would underestimate performance as a protective mechanism to increase opportunity to receive positive feedback. Results from the first model support this prediction but, after receiving positive feedback on the first think paper, these individuals overestimated performance on the second task. We speculate that positive feedback perhaps boosted confidence or reduced fear and anxiety associated with not performing well. As such, these learners were more confident in their performance on their subsequent task. However, average performance on the midterm was, in fact, lower than average performance on the first (and second) think paper. Perhaps these individuals based judgments about their performance on the first task to make predictions for the second task, despite differences in level of difficulty between the tasks. As a result, these individuals were overconfident in their predictions of performance. Finally, it could also be the case that because the tasks differed (a think paper versus a midterm exam), individuals’ calibration differed due to task differences rather than previous feedback. Future research is necessary to disentangle these possibilities.

For performance-approach oriented learners, individuals underestimated performance on the first two tasks but then overestimated performance for the last task. Although performance-approach oriented individuals strive to outperform others, they may “low ball” expectations to be able to interpret feedback positively, that is, “I did better than I expected.” Negative feedback would undermine their sense of ability and contradict their intentions to demonstrate performance. Purposefully gaming positive feedback may be a strategy they use to avoid receiving negative feedback. This conjecture aligns with achievement goal theory. Dweck and Leggett (1988) proposed that individuals with a performance goal orientation might interpret negative feedback as evaluative or judgmental. Accordingly, they may lower their expectations to avoid receiving that
negative feedback. However, for the last task, perhaps as a function of continuous positive feedback, these individuals over-adjusted expectations of their performance, which resulted in an overestimation of their performance for the second think paper.

In contrast to the patterns found for performance-approach oriented learners, individuals high on performance-avoidance orientation overestimated performance on the first two tasks but underestimated performance on the last task. Underestimating performance is consistent with a protective mechanism upon receiving negative feedback (Dweck & Leggett, 1988). However, to interpret the positive relationship between performance-avoidance and calibration bias for the first two tasks, it could be that because these individuals fear failing, they aim for high achievement but their performance is undermined due to mounting worry, concern and fear of failure (Elliot & McGregor, 2001).

4.3. Achievement Goal Regulation

To account for changes in levels of achievement goals, consistent with our predictions, individuals who performed better than they predicted on their first think paper decreased subsequent performance-avoid goal orientation. On the midterm exam and second think paper, individuals who scored lower than they anticipated decreased their performance-approach orientation. No relations were found between feedback and subsequent mastery-approach and mastery-avoid goal orientations. It is tempting to interpret that students in our sample lowered their mastery-approach and mastery-avoid goal orientations because they are poor at estimating their success on these complex and cumulative tasks that require considerable preparation. Because we did not collect data about how students felt about their inaccurate calibration (indeed, no study of which we know has collected such data), we cannot validly speculate this decline results from evolving more realistic perceptions about performance over successive course assignments.

It is also tempting to hypothesize that because the types of tasks students completed changed across time, if the constructs we examined are task-dependent, then patterns might shift because of the type of task rather than as a function of feedback. This is not defensible given that patterns across the same tasks (e.g., the two think papers) were not similar in our study. Moreover, Muis and Edwards (2009) examined the stability of students’ achievement goals over two similar and two different tasks. They found statistically detectable shifts in students’ goals regardless of the comparisons made. Furthermore, we suggest that even if the same tasks were used repetitively, students’ internal representations of those tasks might change after feedback. For example, as Bandura (1997) suggests, when students receive performance feedback on a specific task, their self-efficacy for that same task may change if expectations were not consistent with performance. This, in turn, may alter their self-efficacy for subsequent exposures to that task. Accordingly, internal representations of the “same” task may change and, if previous experience is carried forward, then feedback about performance on the task might change the framework. Without directly asking students about their feedback experiences and perceptions about tasks, the issue of task similarity cannot be resolved. Explanations along these lines must await future research.

4.4. Goal Orientation and Performance

Our path models (Figures 3, 4, and 5) show clearly that goal orientations have varying and sometimes no relations to performance across time, and that calibration bias mediates relations in some cases but not in others. While this could be interpreted as indicating there are no consistent relations between goal orientations, calibration and performance across the measures, we reject this view. Rather, we judge this is an important finding because it allows that goal orientations are more state-like than trait constructs. While Dweck and colleagues (e.g., Dweck & Leggett, 1988; Elliot & Dweck, 1988) conceptualized achievement goals as a personality dimension, our results indicate goals have some instability and may be relatively state-like or task-specific. This is consistent with
recent research on the stability of students’ achievement goals (e.g., Fryer & Elliot, 2007; Senko & Harackiewicz, 2005; Muis & Edwards, 2009).

Although two tasks in our study were similar in nature (i.e., the two think papers) the specific content of the think papers differed, the placement of these tasks in the developmental stream of the course was different, and students may have set different goals for these similar tasks due to differences in content plus accumulated experience. This explanation could help clarify inconsistent relations found in the literature (see Hulleman et al., 2010). Perhaps in each study, relations between each of the tasks and goal orientations were task-specific. This conjecture needs testing in future research as it has significant implications for theorizing about this very frequently studied family of constructs.

Overall, our study indicates that, in an authentic setting where learners experience varied tasks over time, relations among the constructs we examined were not consistent. We submit this is noteworthy because it directs future research to more fully explore why these variations occur across time and context.

4.5. Limitations and Future Directions

We urge researchers to further examine whether tasks, feedback or both change students’ goal orientation framework. While we did not directly examine whether tasks altered the goal framework, future research should specify tasks when measuring students’ goal orientations and compare task-focused goals to general goals across multiple tasks. We recommend that tasks should also vary in levels of similarity although there remains a substantial question of whether similarity should be judged by researchers/instructors or by students themselves. These empirical examinations may help clarify why relations across performance episodes are not consistent within or across studies. We also suggest that future work investigate roles of emotions in achievement goal regulation. Given that emotions such as fear, anxiety and worry are theoretically coupled with avoidance goals and performance-approach goals, it may be that the emotions experienced after feedback is received are more potent predictors of achievement goal regulation (Pekrun, Elliot, & Maier, 2009).

We also recommend larger sample sizes be used when multiple time points are measured to afford using latent growth curve analysis. While our sample is smaller than most, the value added from our research is large. Few studies incorporated a longitudinal component to the study of achievement goals in a real educational context as we did.

Acknowledgement

Support for this research was provided by grants to Krista R. Muis and Philip H. Winne from the Social Sciences and Humanities Research Council of Canada, and the Canada Research Chairs program.

References


~ 35 ~


