Individual vs. Class-Performance Comparison: Impact of Frame of Reference on Students' Outcome Emotions of Pride, Disappointment, Relief, and Shame

Marek Hatala
School of Interactive Arts and Technology
Simon Fraser University
Surrey, British Columbia, Canada
mhatala@sfu.ca

Reyhaneh Ahmadi Nokabadi
School of Interactive Arts and Technology
Simon Fraser University
Surrey, British Columbia, Canada
reyhaneh_ahmadi_nokabadi@sfu.ca

Özge Nilay Yalçın
School of Interactive Arts and Technology
Simon Fraser University
Surrey, British Columbia, Canada
oyalcin@sfu.ca

Abstract

Learning Analytics Dashboards (LADs) make students aware of their progress and outcomes with the goal of inducing reflection and increasing or maintaining motivation. Viewing an outcome-presenting LAD is expected to trigger the causal search process which determines the student's cognitive, emotional and motivational response. Ideally, the information presented in LAD and how it is framed results in a positive emotional and motivational state; however, the knowledge of how to design LADs with desirable impacts is scarce. In a field randomized study with 149 participants, we showed how students' achievement emotions were impacted immediately after viewing a widely deployed LAD in major LMS and a LAD prototype, both of which showed students' assignment outcomes. The students' responses varied between LAD conditions according to students' achievement goal orientation and grade received. This study contributes to the knowledge of how LAD elements and their framing impact students' achievement emotions.

CCS Concepts

• **Human-centered computing → Visualization**; User studies; • **Applied computing → Education**.

Keywords

learning analytics dashboard, achievement emotions, social comparison, individual differences, achievement goal orientation

ACM Reference Format:
Marek Hatala, Reyhaneh Ahmadi Nokabadi, and Özge Nilay Yalçın. 2026. Individual vs. Class-Performance Comparison: Impact of Frame of Reference on Students' Outcome Emotions of Pride, Disappointment, Relief, and Shame. In *LAK26: 16th International Learning Analytics and Knowledge Conference (LAK 2026), April 27–May 01, 2026, Bergen, Norway*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3785022.3785080

1 Introduction

Learning Analytics Dashboards have been deployed and studied as one of the early deployments of Learning Analytics to provide students with means to be aware of their progress, reflect on, motivate them, and change their study behaviours [47]. They represent a feedback mechanism [18]. However, unlike traditional instructor-provided feedback, LADs present information derived from trace data in learning environments. Good feedback, however, is attuned to individual students' needs [18], which creates a challenge for automatically populated LADs. Indeed, the evidence of LADs effectiveness in improving learning outcomes is poor [23, 44], with modest evidence of LADs' positive impact on motivation [23]. Clearly, more research in several directions is needed.

LADs are presented to students with the assumption that they will reflect on the information and respond adaptively. However, LADs are often complex, which may lead to misinterpretation [28], or include elements that paint a conflicting picture [49]. LADs also differ in scope, from overview LADs with many elements [28, 58] to those focusing on a single activity or desirable behaviour [27, 43]. To date, LADs or LA interventions with information elements reflecting concrete learning activities or knowledge have produced findings that are more revealing of how LADs impact students' actions or states (e.g., [7, 27, 43, 49]). Additionally, prior studies have shown that the frame of reference [21, 54], that is, how a student's information is presented alongside other information, is highly influential in directing students' actions or emotions. While peer- or norm-referenced feedback, often implemented following social comparison theory [12], has been found effective [7, 13, 30, 43], other studies report undesirable impacts on students' emotions [49].

LADs do not impact all students in the same way [45]; personal characteristics such as goal orientations [11, 42] can even reverse the effect of the same LAD on learning activities [17, 43]. These findings align with major theories of motivation [10, 31], where goal orientations and achievement emotions are strongly linked [32, 33]. As achievement goal orientation profiles are relatively stable [15], they can provide a basis for LAD personalization.

This study builds on and contributes to research directions identified above. We present experimental evidence from a randomized field experiment with the goal of determining how LADs with precisely defined elements and framing impact students' achievement emotions, which in turn influence students' motivation. Students in two programming courses were presented with the results of their two-week assignments in two different LADs, which differed in their social-comparison frame of reference. We studied how students' emotional responses varied between students with different

LAK 2026, Bergen, Norway

ACM ISBN 979-8-4007-2066-6/26/04
https://doi.org/10.1145/3785022.3785080

achievement goal profiles after viewing LAD, while controlling for students' performance, academic competence, and motivation prior to engaging with the learning activity which outcomes were presented in LADs.

2 Background and Research Questions

2.1 Achievement Emotions and Motivation

Motivational factors have among the largest effect sizes on students' performance [41], reflecting that motivation and learning are mutually supportive [19]. Several motivation theories explain how achievement situations and their outcomes shape motivation. Weiner's Attributional Theory [51] states that when students encounter success or failure, they assign causes that can be characterized by controllability, locus of control, and stability. Combinations of these attributes give rise to achievement emotions, such as pride or guilt for controllable, internal causes (for example effort), or hopelessness for uncontrollable, internal, and stable causes (for example low ability), which in turn influence motivation. Pekrun's Control-Value Theory [31] similarly considers controllability and the value students assign to a task as key determinants of emotional responses. It distinguishes three types of emotions [34]: activity emotions (for example boredom, joy), prospective outcome emotions (hope, hopelessness, anxiety), and retrospective outcome emotions (pride, relief, shame or guilt, disappointment). Expectancy-Value Theory [38] considers task value, expectancy of success, and the cost of engagement as major components of motivation. All three theories explicitly consider students' aptitudes and dispositions, such as Achievement Goal Orientations [42], as moderators of emotional and motivational responses. Students' emotional and motivational states are updated in a cyclical process triggered by engagement in achievement activities and their outcomes.

One of the LADs' goals is to drive motivation [47], and the recent mapping study confirmed moderate success in doing so [23]. Although LADs represent a feedback mechanism (formative or summative) in an achievement setting that triggers emotional and motivational responses, these were rarely studied using the above well-established theoretical frameworks (e.g., [2, 28, 49]), mainly when comparing LA interventions or LADs with a varied frame of reference. Given the outsized impact of motivation on learning outcomes [41], we use achievement emotions, which strongly overlap with motivation [34], to evaluate LADs influence on learners.

2.2 Frames of Reference and Social Comparison

To support students' sensemaking from learning analytics, Wise [54] emphasized the need for a reference frame, that is, a comparison point that may draw on the student's past performance, a teacher-defined standard, or peers' activity or performance. Studies have compared how students respond to these frames with mixed findings (for example, [30, 49]).

Social Comparison (SC) Theory [12] is often cited as a theoretical underpinning for social comparison features in LADs. Different aspects of SC are potentially highly relevant to peer comparison in LA and LADs. The SC is based on the premise of the human drive to compare their performance with others when a standard for comparison is missing. Even when a standard is known, students may not judge success or failure correctly without knowing the performance of others, such as when the exam was very hard, and 100% was not achievable, a situation corresponding to task difficulty in Weiner's theory. The original SC theory and extensive research focused on comparisons with individuals. Indeed, the meta-analysis of 60+ years of research in SC excluded non-individual studies [16].

SC research has mainly examined how individuals select comparators [56]. Such selection depends on comparison goals: for self-evaluation, most people choose upward comparisons even if these evoke negative feelings; for self-improvement, comparison with higher-performing peers is favourable in non-competitive settings; and for self-enhancement, people tend to compare downward to avoid negative affect. In LADs, students have little choice, as they face forced comparisons, presumably with self-evaluation and self-improvement goals in mind. However, even in forced comparisons people are selective about the information they attend to [57], which makes the impact of more complex LADs harder to predict. Several aspects of Festinger's original SC formulation have been expanded or revised. Wood's [56] definition of SC included all socially derived information, such as a group (class) average frequently used in LADs.

Another body of SC research examined the relative strengths of comparison with close individuals as opposed to the cumulative measures. A nameless average allows an individual to make any comparison one desires [36] and a series of experiments [8, 59] showed a strong local dominance effect of comparing with close individuals as opposed to the more remote context like standing in the class. Finally, research on dimensional SC found a social comparison with peers has a significantly larger effect on students' self-evaluation than the dimensional frame (student's performance in other courses) and temporal frame (self-improvement over time) comparisons [55, 60]. Finally, the research has shown that adding information on surrounding dimension(s) (e.g., grades+time spent) can alter the result of the comparison in a way that meets an individual's comparison goals [16, 56].

To summarise, several critical findings in SC research point to the complexity and nuance of the comparison process. More research is needed to understand how elements in LADs and how their values impact learners when placed in various SC framings. Our study contributes to this research.

2.3 Conceptualization of Social Comparison and Differences of Studies in LA and LADs

Social comparison is the most common way [21] to provide a frame of reference [54] in LADs. However, compared to social psychology research on SC above, findings in LA and LAD research on SC's value and usefulness are controversial, which we believe stems from different conceptualization, operationalization, and sources of evidence. Here, we highlight some evidence relevant to our study.

Study in [22] used situated theory of learning and card sorting procedure in focus groups, individual interviews and paired interviews to determine which dataset students would *consider important* in a student-focused LAD. The card sorting revealed the grades, historical comparison data, and online resource use to be among the five most preferred. The authors defined a "Comparison" card as a "Benchmarking comparison against peer classmates: grades, activity, progress." Interestingly, the desire for comparison information was culturally-dependent with international students wanting it strongly, and domestic students (Northern Ireland) being ambivalent to it.

Similarly, using an extensive questionnaire, Divjak et al. [9] asked students about *how important* specific LA or LAD features were for students. The comparison was defined using eight features, such as comparison with generation or course average grades or credits taken, or relation to other students (e.g., among 10% of most successful students). Students rated the comparison features slightly above the midpoint on the unimportant-important scale, behind highly rated features that helped with short-term planning, deadlines, grades, and risk and prediction.

The two above studies captured students' opinions about comparison features in the context of an *envisioned student's LAD*. In contrast, evidence from studies where students actually faced LADs, either in use or hypothetical, paints a more nuanced picture of how social comparison affects emotions, motivation, and performance (e.g., [7, 14, 28, 30]). In Lim et al.'s [28] study, participants were shown LAD variants with *hypothetical* activity completion, study hours, and grades, using four frames of reference: self, course, peers, and course+peers. Using a talk-aloud protocol, the authors found that 69% reported negative affect, including 36% attributing outcomes to lack of effort (which, according to Weiner [51], is a positive attribution) and 30% experiencing negative affect from peer comparison, which triggered anxiety, depression, or stress. Two aspects of this study are critical for interpreting these results: 1) the dashboards displayed very low student engagement and performance, specifically 50% of requirements and about 50% of peers' *average* progress, with study hours declining from 3 to 1 hours weekly and half those of others; and 2) the LADs presented this information after seven weeks, with only three weeks remaining. Despite the negative affect, the LADs were highly motivational (M=5.60/7.00, SD=1.29), with no statistical differences between frames.

While the above study [28] examined students' reactions to the LADs with hypothetical progress, several studies used randomized field experiments to study the impact of SC within actual courses. Brusilovsky et al. [7] deployed MasteryGrid, an open learner model showing students' grasp of course concepts in two variants: comparison with group's *average* and with *group and individual* peers. Students in the individual peers variant significantly outperformed the group variant in 15 measured engagement categories. All students in the individual condition had higher learning gain; however, it was significant only for weak students (i.e., low pre-test scores). Finally, in a detailed usefulness questionnaire administered after *using the system* for the whole semester, students ranked the importance and usefulness of comparison features high, at or above 3.9 on a 5-point scale.

Fleur et al. [14] designed the LAD following Festinger's original theory. The LAD showed nine students with the closest current grade, three performing below and six above the student; the LAD was accessible for the duration of the course. The authors found that the LAD group outperformed the control with a medium effect size, and only 3% of the LAD group failed compared to 18% in the control. The LAD group's extrinsic motivation increased from the first to the last week, while it fell for the control group (N.B., the intrinsic motivation fell for both groups at the same rate).

In our previous qualitative study [1], following the theoretical discussion above, we examined students' attributions, emotions, and motivation for the next assignment after viewing hypothetical LADs showing their current assignment results together with six peers with higher, same, or lower grades. We also examined how additional information altered responses by adding time spent on the assignment (higher or lower) or students' ability (higher or lower). Students' achievement goal orientation profiles moderated these reactions; the three clusters were: Cluster 1, high mastery and low performance-avoidance; Cluster 2, high mastery and high performance-avoidance; and Cluster 3, low mastery and low performance-approach. The LAD showing upward comparison where peers had higher grades and spent more time triggered desirable attributions to low effort [51] in 100%, 80%, and 71.4% of students in Clusters 1–3, and was motivating for 75%, 80%, and 57%. The most demotivating LAD differed by cluster: same grades with higher time for Cluster 1 (50% demotivated), and higher grades with lower time for Clusters 2 and 3 (60% and 86%).

2.4 Present Study

As previous sections showed, social comparison in LADs can motivate students or positively affect their behaviours when deployed in courses or examined in hypothetical scenarios, even if SC features are not among the most preferred. However, the reviewed studies also show that LADs can generate diverse attributional, emotional, and motivational responses. Many authors have therefore advocated for further research to understand these impacts and personalize LAD content and presentation [5, 28, 45, 49]. We aim to address two open questions. First, SC research argues that comparing with individuals can elicit stronger responses than comparing with group measures such as averages. This was observed in [7], and other studies also reported strong effects of individual comparison. Yet most dashboards use group averages as the comparator. In this study, we compare the impacts of two LADs: one with individual comparison and one with group comparison. Secondly, several studies [1, 43] show that achievement goal orientation moderates students' responses to LADs and is a promising basis for personalization. We therefore investigate how emotional impacts differ across students' achievement goal orientations. We formulate our research questions as follows:

RQ1: What are the emotional impacts of LADs that use social comparison with individuals versus class performance measures?

RQ2: How do the emotional impacts of these two social comparison frames of reference differ for students with different achievement goal orientation profiles?

3 Method

We answer our research questions using a randomized field study, comparing the impact of the most motivational dashboard from our prior research [1] (higher grades - higher time spent) on students achievement emotions and motivation. As a control baseline, we compare the dashboard to the one embedded in the Gradebook of Canvas[1] LMS, which uses class performance distribution.

[1]https://www.instructure.com/canvas

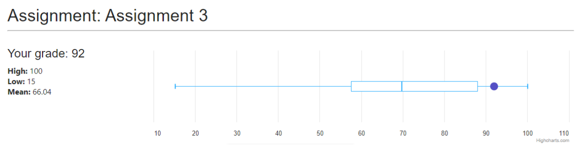

Figure 1: Class Performance dashboard (CP-D) inspired by Canvas - the blue circle on the boxplot shows the viewer's grade.

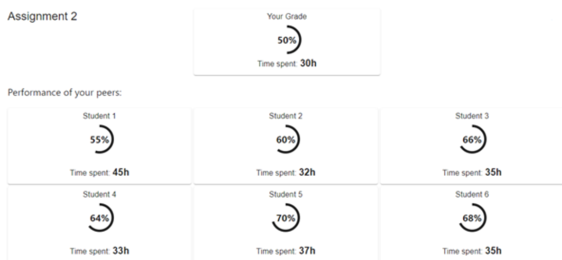

Figure 2: Individual Comparison dashboard (IC-D) shows the student and six higher performing peers.

3.1 Dashboards Design and Implementation

Figure 1 shows the dashboard for the class-performance comparison (CP-D), showing the boxplot of student grades. The dashboard design mimics the dashboard embedded in the Gradebook tool of Canvas, the leading LMS in North America (50% of student enrollments, [20]). As such, it can be considered a 'control' condition, as all students using Canvas are exposed to this type of dashboard.

The individual comparison dashboard (IC-D) in Figure 2 shows six peers with higher scores and higher time spent on the assignment. As described in Section 2.3, this dashboard was the most motivational dashboard for all students in our prior study, regardless of their achievement goals profile. We chose the peers who got 5 to 20 percent higher grades than the viewer and reported spending more time. If such students did not exist in the class, we generated students with fictitious grades and time spent[2].

The System Usability Scale (SUS) [6] administered at the end of the study showed both LADs as highly usable. CP-D score was 73.25, with a high learnability component (71.25), likely due to familiarity with the Canvas dashboard. IC-D scored 69.88 overall, with a learnability score of 62.80.

3.2 Instruments

The demographics questionnaire collected information such as age, gender, prerequisite programming course grade, GPA, and number of courses taken in the semester. Students' achievement goal orientations were measured using the mastery, performance-approach, and performance-avoidance subscales of the Patterns of Adaptive Learning Scales (PALS) [29], each rated on a 5-point Likert scale. Because the subscales contained different numbers of items, we used the average item score to obtain values on the common 1–5 range.

Achievement emotions were measured with the Revised Achievement Emotions Questionnaire (AEQ-R) [35], which includes components for emotions before, during, and after a learning activity. This study analyzes four retrospective outcome emotions: pride, disappointment, shame, and relief. As with PALS, we used average item scores to place each emotion on the 1–5 scale. These emotions were captured immediately after seeing assignment results in LADs.

Finally, we measured three motivational constructs following the expectancy-value-cost model of motivation [3] using the questionnaire from [25]. The expectancy dimension consists of students' ability beliefs and expectancy beliefs. The value dimension captures three positive contributors to the value of a task: intrinsic value, utility value, and attainment value. The cost dimension refers to negative aspects of engaging in an activity, such as the effort and time required for success and the potential for negative psychological states, such as fear of failure. We replaced the three direct questions measuring values with the six-item Task Value subscale from the MSLQ [37], as these items cover course, content area, and subject matter. Motivation constructs were measured on a 7-point Likert scale; for each construct, the average item score was used to obtain values on the 1–7 range.

3.3 Procedure

After signing the consent form, demographic and PALS questionnaires were sent to the participants recruited from two programming courses in the first and second years. The courses used the same structure with five programming assignments with 2-3 weeks due dates. We decided to utilize the dashboards for the grades of the second and third assignments in both classes, allowing students to adjust to the course requirements. After each assignment was released on Canvas, AEQ-before and EVC questionnaires were sent to the participants. The next questionnaire, AEQ-during, was sent to them 3 to 4 days before the deadline. Students reported their time spent as part of the assignment submission. The assignments were marked promptly, and study participants were notified that their grades were available for viewing in the dashboard. They were asked to follow the link to the LADs online according to the treatment conditions. The final questionnaire, AEQ-after, was linked to the "Take me to the questionnaires" button in the dashboard. After clicking on the button, the viewer was able to complete the said questionnaire. For the second assignment used in the study (A3 in the courses), the questionnaire regarding achievement emotions and motivation before the next assignment was attached to the end of the last questionnaire for the current assignment. The SUS questionnaire was collected at the end of the study. All questionnaires were created online on Survey Monkey[3].

3.4 Participants

157 participants were recruited from the first and second-year programming courses in the interdisciplinary program at the Canadian research-intensive university. The University Research Ethics Board approved the study; the participants received a 2% participation

[2] The participants were debriefed about the deception at the study end and were given an opportunity to withdraw their data.

[3] www.surveymonkey.com

Table 1: Descriptive statistics for variables of interest (mean and SD values)

	CP-D Group	IC-D Group	Cluster 1	Cluster 2	Cluster 3
n	98	51	51	29	68
GPA[a]	3.16 (0.41)	3.13 (0.58)	3.18 (0.62)	3.25 (0.57)	3.05 (0.42)
Prerequisite Grade[a]	79.58 (9.78)	79.00 (10.33)	80.92 (9.61)	81.43 (8.48)	77.15 (10.75)
Value[a,b]	4.88 (1.31)	5.03 (1.34)	5.50 (0.91)	5.39 (1.44)	4.44 (1.35)
Expectancy[a,b]	4.68 (1.28)	4.86 (1.23)	5.17 (1.15)	5.27 (1.13)	4.34 (1.22)
Cost[a]	4.53 (1.28)	4.35 (1.17)	4.38 (1.18)	4.07 (1.34)	4.58 (1.16)
Mastery[c]	-	-	3.47 (0.60)	3.87 (0.41)	2.92 (0.74)
Performance-approach[d]	-	-	4.00 (0.56)	2.72 (0.81)	2.66 (0.69)
Performance-avoidance[c]	-	-	3.52 (0.74)	1.70 (0.44)	2.77 (0.58)

[a] No significant differences between dashboard groups.
[b] No significant differences between clusters.
[c] All clusters significantly different.
[d] Cluster 1 significantly higher than clusters 2 and 3.

credit towards their final grade. 157 participants filled out the demographics questionnaire: 144 were 18-24 old, 13 were 25-30; 88 were women, 62 men, three non-binary, and four did not disclose. The majority (72) took three courses in the semester, 47 more than 3, and 38 less than three. Participants were sorted within each course by their prerequisite course grade and alternately assigned to the CP-D or IC-D conditions in an approximately 1:2 ratio, as biofeedback was collected from the IC-D group (partly analyzed here [48]). Two students withdrew from the course before the first study assignment, one withdrew consent at the end, and five discontinued participation; all such data were removed. Table 1 shows descriptive statistics for the remaining 149 participants. The two LAD groups were balanced in the overall sample: t-tests indicated no significant differences in academic performance (GPA, prerequisite grade) or initial motivation constructs (value, expectancy, cost).

To support person-centred analysis [46] along students' achievement goals, the three PALS subscale values were used to cluster 148 participants[4] using agglomerative clustering (`hclust` in R) with Ward's method. The subscale values were normalized before clustering, and the Euclidean metric was used. After examining the dendrogram and the Silhouette statistic [39], we selected a three-cluster solution; the clusters are summarized in Table 1. Across all three clusters, t-tests indicated no significant differences between the CP-D and IC-D conditions on any variables in Table 1, suggesting that random assignment did not bias cluster composition. We labelled and interpreted the clusters as follows:

- *Cluster 1 – high performance-approach, medium mastery, and medium performance-avoidance:* These students care most about outperforming others, while also seeking self-improvement and avoiding being seen as incompetent.
- *Cluster 2 – high mastery, medium performance-approach, and low performance-avoidance:* These students care about mastering the course content and achieving good grades, not worrying about being seen as incompetent.
- *Cluster 3 – medium mastery, low performance-approach, and medium performance-avoidance:* These students show moderate interest in mastering the content, do not care about grades, and show little concern about incompetece.

[4]One participant had missing PALS data and was not included in the cluster analysis.

3.5 Statistical Analyses

We have run statistical analyses on the pooled data from five offerings of two courses. In each course, students viewed their grades in either the Class-performance dashboard (CP-D) or Individual comparison dashboard (IC-D) after completing each assignment. Since the assignments were nested within two courses, we have used hierarchical linear mixed effect models (HLME) [40] utilizing `lmerTest` R package [26]. Since we have captured students' emotions and motivational constructs for two consecutive assignments, we have accounted for repeated measures from students. We have fitted a separate model for each retrospective outcome emotion after viewing the dashboard.

The dashboards' main purpose is to inform students about the outcome of their work on the assignment. According to Weiner [51], the student's grade communicates success or failure, instigating the cause that will determine the student's emotional and motivational state. According to Control-value theory [31] and Expectancy-value theory [10], students' emotional and motivational reactions are reciprocal to the achievement outcomes. Grades are interpreted as success or failure based on several factors. We have used GPA as it represented academic capability; we have used value, expectancy, and cost, measured prior to the student's work on the assignment, to represent the student's motivation state in the appraisal step of the cyclical process linking emotions, motivations and effects [31, p.328]. To capture the impact of the LAD type, i.e., how the student's grade is presented, we included a categorical term for the LAD seen (0 for CP-D, and 1 for IC-D). Finally, the interaction term Grade x Dashboard allowed us to capture the varying impact of LAD type across the range of grades received.

To answer RQ2, we fitted the same model separately for each AGO cluster. This choice simplifies interpretation and aligns with an envisioned deployment scenario where a student is first assigned to a cluster and the corresponding model is then used to estimate the emotional impact of the two LADs.

4 Results

The outcome variables in the models were tested for normality using the Shappiro-Wilks test and by examining skew and kurtosis. Violations of the normality for emotion and motivation measures were minor, and no transformations were applied. The predictor

Table 2: Mixed effect models for retrospective outcome emotions for the whole sample and clusters 1-3. Estimates and standard errors shown for fixed effects. F-statistics for the IC-D condition and Grade × IC-D interaction are reported at the bottom of the table.

	Pride				**Relief**			
	All	Cluster 1	Cluster 2	Cluster 3	All	Cluster 1	Cluster 2	Cluster 3
Intercept	-0.19 (0.58)	0.93 (0.90)	-0.09 (1.18)	-0.70 (0.90)	1.00 (0.58)†	1.24 (1.10)	-0.64 (1.12)	0.97 (0.82)
GPA	-0.20 (0.11)†	-0.24 (0.14) †	-0.46 (0.25) †	-0.03 (0.20)	0.00 (0.11)	-0.16 (0.17)	0.09 (0.23)	0.09 (0.18)
Value	0.09 (0.05)†	0.07 (0.10)	0.10 (0.10)	0.02 (0.07)	0.12 (0.05)*	0.12 (0.12)	0.19 (0.10)†	0.09 (0.06)
Expectancy	0.04 (0.05)	-0.04 (0.08)	0.01 (0.11)	0.11 (0.07)	-0.09 (0.05)†	0.11 (0.10)	0.18 (0.12)	-0.13 (0.07)*
Cost	0.03 (0.05)	-0.04 (0.08)	0.01 (0.11)	0.09 (0.07)	0.11 (0.05)*	0.11 (0.10)	0.00 (0.10)	0.13 (0.06)*
Grade	0.04 (0.00)***	0.04 (0.01)***	0.04 (0.01)***	0.04 (0.01)***	0.02 (0.00)***	0.03 (0.01)***	0.02 (0.01)*	0.02 (0.01)***
IC-D	0.97 (0.46)*	-0.69 (0.72)	2.47 (1.10)*	1.04 (0.66)	0.93 (0.45)*	0.60 (0.89)	2.8 (1.04)*	1.02 (0.59)†
Grade x IC-D	-0.01 (0.01)*	0.00 (0.01)	-0.03 (0.01)*	-0.01 (0.01)†	-0.01 (0.01)†	-0.01 (0.01)	-0.03 (0.01)*	-0.01 (0.01)
Marginal R^2	0.41	0.49	0.48	0.34	0.18	0.22	0.41	0.20
Conditional R^2	0.44	0.71	0.48	0.44	0.38	0.42	0.41	0.40
F (IC-D)	F(1, 178)=4.6*	F(1, 73)=0.9	F(1, 40)=5.1*	F(1, 80)=2.5	F(1, 178)=5.9*	F(1, 70)=0.6	F(1, 36)=9.6*	F(1, 109)=2.4
F (Grade x IC-D)	F(1, 183)=5.0*	F(1, 73)=0.3	F(1, 40)=4.3*	F(1, 84)=3.4†	F(1, 185)=5.1*	F(1, 74)=0.6	F(1, 37)=8.4*	F(1, 110)=2.3

	Disappointment				**Shame**			
	All	Cluster 1	Cluster 2	Cluster 3	All	Cluster 1	Cluster 2	Cluster 3
Intercept	6.62 (0.72)***	5.49 (1.22)***	6.52 (1.21)***	6.90 (1.13)***	5.52 (0.74)***	4.44 (1.27)***	4.31 (1.38)*	6.47 (0.99)***
GPA	-0.03 (0.14)	0.01 (0.19)	0.38 (0.24)	-0.30 (0.25)	-0.06 (0.15)	0.01 (0.22)	0.49 (0.32)	-0.43 (0.22)†
Value	0.09 (0.06)	0.15 (0.13)	0.18 (0.10)†	0.15 (0.09)†	0.04 (0.06)	0.24 (0.12)*	0.25 (0.12)*	0.06 (0.08)
Expectancy	-0.25 (0.06)***	-0.11 (0.10)	-0.26 (0.12)*	-0.31 (0.09)***	-0.28 (0.06)***	-0.32 (0.10)*	-0.14 (0.14)	-0.37 (0.08)***
Cost	-0.06 (0.06)	0.12 (0.10)	-0.03 (0.10)	-0.14 (0.09)	-0.01 (0.06)	0.16 (0.10)	0.12 (0.11)	-0.16 (0.08)*
Grade	-0.03 (0.01)***	-0.04 (0.01)***	-0.06 (0.01)***	-0.02 (0.01)*	-0.02 (0.01)***	-0.03 (0.01)*	-0.05 (0.01)***	0.00 (0.01)
IC-D	-1.37 (0.56)*	-0.74 (0.96)	-3.17 (1.02)*	-1.27 (0.82)	-1.08 (0.57)†	-1.34 (0.97)	-4.24 (1.04)***	-0.33 (0.72)
Grade x IC-D	0.02 (0.01)*	0.01 (0.01)	0.03 (0.01)*	0.02 (0.01)	0.01 (0.01)†	0.02 (0.01)	0.04 (0.01)***	0.00 (0.01)
Marginal R^2	0.27	0.45	0.51	0.18	0.18	0.25	0.36	0.21
Conditional R^2	0.32	0.58	0.68	0.21	0.54	0.76	0.89	0.29
F (IC-D)	F(1, 184)=4.4*	F(1, 64)=0.5	F(1, 40)=7.3*	F(1, 75)=3.0†	F(1, 219)=3.6†	F(1, 77)=1.9	F(1,34)=16.5***	F(1, 79)=0.2
F (Grade xICD)	F(1, 193)=3.6†	F(1, 68)=0.5	F(1, 40)=7.4*	F(1, 78)=2.1	F(1, 226)=3.2†	F(1, 72)=1.9	F(1,29)=14.8***	F(1, 82)=0.1

p-values: † p<.1, * p<.05, ** p<.01, *** p<.001

variable assignment grade was left-skewed (-0.86), which is below 2, which signifies normality [52]. The fit of HLME models is reported using two values of R^2. The *marginal R^2* represents variance explained by predictor variables alone; the *conditional R^2* captures all variance explained by the model, both from predictors and variance attributed to the groups within the environment, i.e. course, assignment, and individual students (because of repeated measurements).

Table 2 reports R^2 values for fitted models for all students and for the models fitted for each cluster. Except two cases, the models fitted for each cluster showed a better fit than the model fitted to the full sample; especially Cluster 1 and 2 models for pride, relief and disappointment. The grade was the most significant predictor of each emotion in every model, except for shame in Cluster 3. GPA was not a significant predictor. Value was a significant predictor for shame for Clusters 1 and 2; and for whole sample for relief. Expectancy was a significant predictor in most models for disappointment and shame. Cost was significant predictor for relief for whole sample and Cluster 3.

4.1 Grade x Dashboard Interaction Effects

The main purpose of this study was to examine the effect of two frames of reference on students' retrospective outcome emotions after viewing dashboards with different types of social comparison. Because hierarchical linear mixed-effect models with interaction terms are difficult to interpret, we visualized the interaction between Grade and LADs using the R `effects`[5] package. Figure 3 presents the modeled values for emotions and motivational constructs for all participants and for each cluster. The error bars at different points reflect the influence of covariates (GPA and the motivational constructs of value, expectancy, and cost measured before the assignment). The dashboard condition, that is, the dashboard itself or the Grade × Dashboard interaction, was a significant predictor in the overall sample for pride, relief, and disappointment, and marginal for shame. It was not a significant predictor of any emotions in Cluster 1, but it was a significant predictor of all emotions in Cluster 2. For Cluster 3, the dashboard condition was marginally significant for pride and relief. In Figure 3, significance is typically indicated by crossover of prediction lines or minimal overlap of error bars. Full F-statistics and p-values for dashboard condition and Grade × Dashboard interaction terms are included in Table 2.

5 Discussion

In this section, we first discuss the findings concerning shared observations between models before examining the individual retrospective outcome emotions.

[5]https://cran.r-project.org/package=effects

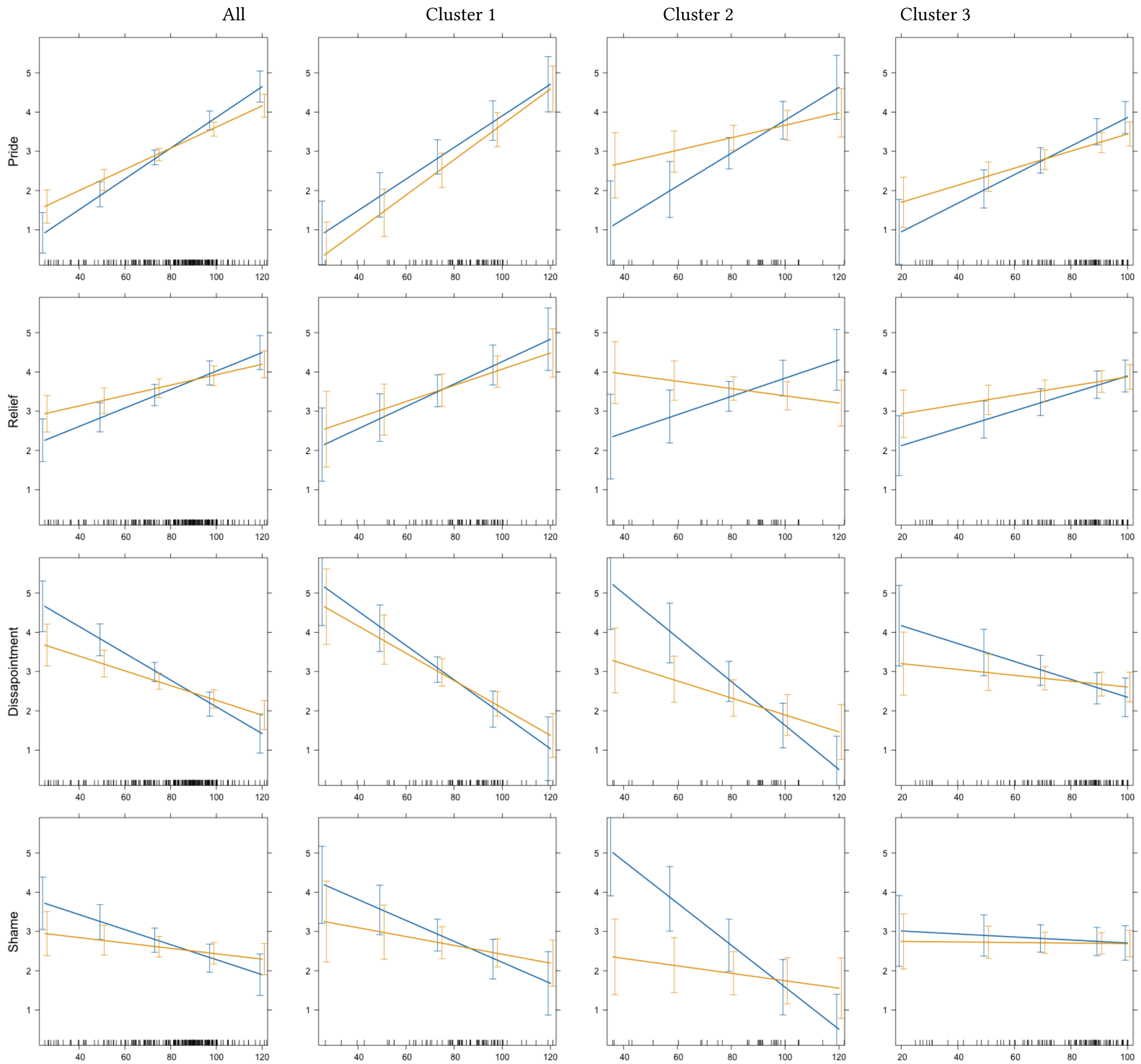

Figure 3: Retrospective outcome emotions modeled values for CP-D (blue) and IC-D (orange). Horizontal axes: assignment grades from 25% (left) to 100% (right), linear scale, inside ticks: individual students marks.

5.1 Overall Model Characteristics

As noted in the previous section, the close values of marginal and conditional R^2 indicate that the course and assignments (hierarchical components) had minimal impact on the emotional effect of grades presented in the dashboards. A few exceptions emerged for shame, mainly in Clusters 1 and 2. For Cluster 1, 55.0% of variance was within students and 8.6% at the assignment level; for Cluster 2, the respective values were 51.7% and 17.8%. Students in both clusters share a higher mastery orientation (see Table 1), and it appears that their experience of shame was more strongly shaped by the assignment context. In other words, they responded with shame differently depending on how they perceived the assignment. We revisit this point in the discussion below. Secondly, as expected, grade strongly and significantly predicted all emotions. Theories of emotion [10, 31, 51] indicate that achievement outcomes trigger emotions in different ways. Because grade reflects the range from failure to success, it was a positive predictor of pride and relief and a negative predictor of disappointment and shame. A few exceptions can be seen in Figure 3, where lines have a low slope. These occur mostly for shame in the individual comparison LAD (IC-D), suggesting that IC-D has the potential to soften shame even among very low-performing students. We also observed that the LAD condition and its interaction with grade were strong predictors for several emotions and clusters, underscoring the varying impact of LAD conditions on students' emotional experiences.

The most striking observation from Figure 3 is that the slope of the lines predicting retrospective outcome emotions after viewing IC-D is flatter than the one using CP-D (i.e., Canvas embedded default), often with the observed crossover point between 75 and 85% grade for all measured emotions. Considering that the average and median grade for the assignments were 78% and 85% respectively, this crossover corresponds to the students viewing themselves above these values in CP-D (see Figure 1). In general, the IC-D had a more positive impact on achievement emotions for students below the crossover point: higher pride, lower disappointment, lower shame, and higher relief. Please note that these findings were reversed for Cluster 1 for pride.

It is especially critical that in IC-D, the low-performing students feel more pride and less disappointment and shame. Pride is a positive activating emotion [34]; authentic pride has been shown to drive behavioural changes for low performers [50] and incentivizes perseverance [53]. For the students who achieved higher grades, they see themselves at the far right of the CP-D above other students in the class. However, lower pride for high-grade students in the IC-D may be to some extent the result of how we populated the dashboard with peers with even higher grades (often created synthetically), as pointed out in Section 5.3. Cluster 1 groups students with exceptionally high performance-approach orientation score, 4.0 out of 5, and also high performance-avoidance score (3.52 of 5). We can observe generally steeper slopes for pride and other emotions in this cluster, driven by grade value itself, with minimal differences between the LAD conditions. This was confirmed by non-significant coefficients for all four emotions for Cluster 1.

Disappointment is a negative deactivating emotion that "relates to failure when success has been expected" [34]. This is confirmed by the models, where expectancy before the assignment was a significant negative predictor of disappointment (except in Cluster 1, where grade was the main driver, as discussed above). At the lower end of the grade scale, the measured disappointment aligns with this definition. In contrast, for students with high scores, disappointment reflects the sense that their already strong grades still fell short of their expectations. While both LADs show students low grades that objectively represent failure in terms of passing the course, the individual comparison dashboard (IC-D) also shows six peers with somewhat higher grades and more time spent, providing a rationale for the low mark. As discussed above, IC-D softened the experience of disappointment for low-performing students, particularly in Cluster 2 (significant predictors) and to some extent in Cluster 3. Students in Cluster 2 had the highest mastery orientation (3.87 of 5), which aligns with the explanation that they interpreted IC-D as offering a rationale for their poor results. This leads to lower disappointment (and shame) than when viewing the class-performance dashboard (CP-D), for which the slope is particularly steep, driven by received grade.

Shame is typically linked to internal and uncontrollable causes, such as low ability [31, 51]. Expectancy was a very strong negative predictor of shame, even stronger than for disappointment, indicating that students tended to direct their feelings toward themselves. For students with low grades, the CP-D displayed most peers above them, which naturally could trigger feelings of inadequacy. The IC-D, however, countered this effect by providing an explanation for the low grade in the form of lower time spent, resulting in an almost flat curve for shame across the grade range. Interestingly, for Cluster 3, which had extremely low performance-avoidance (1.70 of 5), grade had only a minimal effect and there were no observable differences between dashboards.

Relief is a positive deactivating emotion, opposite in meaning to disappointment: it occurs after success when failure was expected [34]. Expectancy was a small but significant negative predictor overall and for Cluster 3. Cost was a significant positive predictor, indicating greater relief among students who perceived the assignment as requiring more time and effort (significant overall and for Cluster 3). As shown in Table 2, the dashboard type and its interaction with grade were significant predictors overall and for Cluster 2, the cluster characterized by high performance-approach orientation. Figure 3 further shows that, for Cluster 2, the slope for relief was negative for IC-D, indicating that this dashboard provided greater relief for low-performing students.

5.2 Achievement Goals Cluster Differences

Due to space limits, we focus only on differences between clusters in the impact of LADs. The clusters represent students' achievement orientation profiles. Clusters 1 and 2 had relatively high mastery orientation, while Cluster 3 was just below neutral. The largest differences between clusters were in their performance orientations. Cluster 1 was highest in both performance-approach and performance-avoidance, whereas Cluster 2 had lower performance-avoidance than Cluster 3, and both were significantly lower than Cluster 1.

Cluster 1 achievement goals were high across all three orientations. The strong performance-approach orientation, focused on absolute performance as represented by grade, was reflected in steep slopes for pride and disappointment with respect to grade. The impact of LAD type was observed only for shame, where the individual comparison dashboard softened the experience of shame for low-performing students, though this effect was not significant. Overall, our results indicate that high performance approach students' retrospective outcome emotions were driven mainly by the absolute value of their grade and were not influenced by how the outcome was framed in the two LAD conditions.

Cluster 2 had a very high mastery orientation and very low performance-avoidance orientation. This combination is indicative of a strong focus on mastering the content, with less concern for how competence is perceived by others. Interestingly, the dashboard type and its interaction with grade appeared to have a notable impact on these students' retrospective outcome emotions. Among students who viewed the class-performance dashboard, emotions varied widely across the grade range: pride increased from about 1 to 5, disappointment and shame decreased from about 5 to 1, and relief stayed between 2.5 and 4. In contrast, in the individual comparison dashboard, pride ranged only from 2.5 to 3.8, disappointment from 3.3 to 1.5, and shame remained below neutral (2.4 to 1.8). For relief, as discussed above, the slope reversed relative to the class-performance dashboard, ranging from 4 to 3.2. These results suggest that for high-mastery students in this cluster, particularly when they perform below the class average, the class-performance dashboard may elicit a more negative emotional profile. Despite lower pride for higher grades, there is an indication that such students could benefit from being presented with individual comparison dashboards.

Finally, Cluster 3 represents students with the lowest mastery orientation and relatively low performance orientations. The impact of LAD type on these students followed a pattern similar to Cluster 2 for pride and disappointment, but the effects were weaker and not significant within the ranges of their responses. These patterns are consistent with students' ambivalence toward the course content and their performance. Students reported only weak relief at the high end of the grade scale and remained largely neutral with respect to shame, regardless of LAD type. Given the slightly more positive emotional profile observed after viewing the individual comparison dashboard, these students may benefit more from this format.

5.3 Limitations

This randomized control study is executed within a specific context, our findings are specific to an interdisciplinary program's first and second-year programming courses. Such courses attracted students with varied interests, bringing forward their motivational differences. Repeating the study in courses with a more homogeneous population in terms of their motivation toward the course subjects would complement our results. The second limitation relates to the IC-D dashboard. Generally, it may be difficult for students to estimate the time spent on an assignment. However, working in a special programming environment represents a distinct activity with a higher chance of close time estimation compared to other assessment activities, such as estimating the amount of time spent on an essay, which may involve a period of conceptualizing and research in opportune times over a more extended period. Thirdly, our rigid method of selecting peers for comparison may have affected the emotions and motivation of students with high grades. Our findings show that the CP-D generally had more positive outcomes for high-performing students. In IC-D, all students saw higher performing peers getting better grades and spending more time, triggering negative reactions for high-achieving students, such as those reported in [4]. If individual comparison is to be considered for the students, as recommended for Cluster 2 and 3, a better strategy is needed for the high-performing students.

5.4 Implications for Research and Practice

In alignment with Control-Value Theory [31] and Expectancy-Value Theory [10], students' achievement goal orientation profiles were critical in determining which LAD was more suitable, leading to (speculative) practical recommendations for LAD choice for individual students. Our findings highlight the importance of studying dashboards at the granularity of the information elements they contain, as well as their framing. The IC-D used in this study was selected as the most motivational for all AGO profiles after a prior rigorous evaluation of 13 dashboard variants [1]. The CP-D was a replica of the LMS Canvas dashboard used by millions of students, which we did not alter to allow comparison with a natural control condition. A direction for future work is to include aggregate time-spent information in the CP-D. The broader research agenda should examine the impact of LAD elements easily derived from LMS trace data (for example file access [24, 28]). Another direction is studying how multiple LAD elements, possibly showing conflicting information, affect students' emotions and motivation (for example [2]). Further, research should examine both how long elicited emotions persist and whether they lead to changes in students' behaviours, as we have done here [61].

Studies examining students' emotions and motivation in LADs and LA at this level of detail are scarce, making comparisons with existing work difficult. We described the study by Lim et al. [28] in Section 2.3, where peer comparison led to negative affect but positive motivational impacts for hypothetical very low-performing students. A recent experimental field study by Weidlich et al. [49] investigated the impact of personalized feedback in four reference frames: feedback-as-usual, peer comparison implemented as class rank, course expectations referenced feedback, and a combined peer and course frame. Although students found the latter three frames useful, the authors concluded that social comparison feedback was "comparably detrimental to motivational regulation" (p. 13). However, peer-referenced feedback did not statistically affect motivation, while the other two frames did. Peer comparison slightly lowered pride and increased shame, but neither was statistically significant. The impact of performance information was similar to ours. Study [49] and ours share similarities; rank comparison in the peer-referenced frame is comparable to our class-performance dashboard, where students see their position in class via quartiles.

Practical implications. This study examined only a few LAD elements, limiting its practical impact. Nevertheless, findings regarding the CP-D dashboard have practical value, as a similar dashboard is included in Canvas. Our results indicate that such a dashboard is detrimental to low-performing students' emotions and should be carefully reexamined.

6 Conclusions

This experimental field study examined how two dashboards, using different peer comparison methods, affect students' achievement emotions and motivation. While the grade displayed was the primary driver of emotional responses, we also found a significant impact of dashboard type on the retrospective outcome emotions of pride, relief, disappointment, and shame. Our results suggest that only high-performing students may benefit from a dashboard showing summarized class performance, whereas students with grades below the median may benefit more from a dashboard presenting a small number of slightly higher performing peers. Among students performing below the class average, those with high mastery and low performance goal orientations appeared most responsive to the dashboard type, showing the greatest benefit from the individual comparison format.

Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada, RGPIN-2018-06071.

References

[1] Kimia Aghaei. 2023. *Investigating Students' Emotional and Motivational Responses to Dashboards with Varying Social Comparison Groups.* MSc Thesis. Simon Fraser University. https://summit.sfu.ca/item/36716

[2] Kimia Aghaei, Marek Hatala, and Alireza Mogharrab. 2023. How Students' Emotion and Motivation Changes after Viewing Dashboards with Varied Social Comparison Group: A Qualitative Study. In *The 13th Int. Conf. on Learning Analytics & Knowledge.* 663–669. https://doi.org/10.1145/3576050.3576107

[3] Kenneth E. Barron and Chris S. Hulleman. 2015. Expectancy-Value-Cost Model of Motivation. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* 8, 2006 (2015), 503–509. https://doi.org/10.1016/B978-0-08-097086-8.26099-6

[4] Liz Bennett and Sue Folley. 2021. Students' Emotional Reactions to Social Comparison via a Learner Dashboard. In *Visualizations and Dashboards for Learning Analytics*, M. Sahin and D. Ifenthaler (Eds.). Springer, Cham, 233–249. https://doi.org/10.1007/978-3-030-81222-5{_}11

[5] Matthew L. Bernacki, Meghan J. Greene, and Nikki G. Lobczowski. 2021. A Systematic Review of Research on Personalized Learning: Personalized by Whom, to What, How, and for What Purpose(s)? *Educational Psychology Review* 33 (2021), 1675–1715. https://doi.org/10.1007/s10648-021-09615-8

[6] John Brooke. 1996. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation In Industry*, Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Bernard Weerdmeester (Eds.). 207–212. https://doi.org/10.1201/9781498710411-35

[7] Peter Brusilovsky, Sibel Somyurek, Julio Guerra, Roya Hosseini, Vladimir Zadorozhny, and Paula J. Durlach. 2016. Open Social Student Modeling for Personalized Learning. *IEEE Transactions on Emerging Topics in Computing* 4, 3 (2016), 450–461. https://doi.org/10.1109/TETC.2015.2501243

[8] Justin T. Buckingham and Mark D. Alicke. 2002. The influence of individual versus aggregate social comparison and the presence of others on self-evaluations. *Journal of Personality and Social Psychology* 83, 5 (2002), 1117–1130. https://doi.org/10.1037/0022-3514.83.5.1117

[9] Blazenka Divjak, Barbi Svetec, and Damir Horvat. 2023. Learning analytics dashboards: What do students actually ask for? *ACM International Conference Proceeding Series* (2023), 44–56. https://doi.org/10.1145/3576050.3576141

[10] Jacquelynne S. Eccles and Allan Wigfield. 2020. From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology* 61, May (2020). https://doi.org/10.1016/j.cedpsych.2020.101859

[11] Andrew J. Elliot, Kou Murayama, and Reinhard Pekrun. 2011. A 3 × 2 achievement goal model. *Journal of Educational Psychology* 103, 3 (2011), 632–648. https://doi.org/10.1037/a0023952

[12] Leon Festinger. 1954. A Theory of Social Comparison Processes. *Human Relations* 7, 2 (1954), 117–140. https://doi.org/10.1177/001872675400700202

[13] Damien S. Fleur, Wouter van den Bos, and Bert Bredeweg. 2020. *Learning analytics dashboard for motivation and performance.* Vol. 12149 LNCS. Springer International Publishing. 411–419 pages. https://doi.org/10.1007/978-3-030-49663-0{_}51

[14] Damien S. Fleur, Wouter van den Bos, and Bert Bredeweg. 2023. Social comparison in learning analytics dashboard supporting motivation and academic achievement. *Computers and Education Open* 4, February (2023), 100130. https://doi.org/10.1016/j.caeo.2023.100130

[15] James W. Fryer and Andrew J. Elliot. 2007. Stability and Change in Achievement Goals. *Journal of Educational Psychology* 99, 4 (2007), 700–714. https://doi.org/10.1037/0022-0663.99.4.700

[16] J. P. Gerber, Ladd Wheeler, and Jerry Suls. 2018. A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin* 144, 2 (2018), 177–197. https://doi.org/10.1037/bul0000127

[17] Marek Hatala and Sanam Shirazi Beheshitha. 2016. Associations Between Students ' Approaches to Learning and Learning Analytics Visualizations. In *Proceedings of the LAK 2016 Workshop on Learning Analytics for Learners*, Susan Bull, Blandine M. Ginon, Judy Kay, Michael D. Kickmeier-Rust, and Matthew D. Johnson (Eds.), Vol. 1596. CEUR-WS, Edinburgh, Scotland, 3–10. http://ceur-ws.org/Vol-1596/paper1.pdf

[18] John Hattie and Helen Timperley. 2007. The Power of Feedback. *Review of Educational Research* 77, 1 (2007), 81–112. https://doi.org/10.3102/003465430298487

[19] Suzanne E. Hidi and K. Ann Renninger. 2019. Introduction: Motivation and its relation to learning. *The Cambridge Handbook of Motivation and Learning* (2019), 1–12.

[20] Phil Hill. 2025. State of Higher Ed LMS Market for US and Canada: Year-End 2024 Edition. *On EdTech Newsletter* (2025). https://onedtech.philhillaa.com/p/state-of-lms-market-us-canada-year-end-2024

[21] Ioana Jivet, Maren Scheffel, Marcus Specht, and Hendrik Drachsler. 2018. License to evaluate: Preparing learning analytics dashboards for educational practice. *ACM International Conference Proceeding Series* (2018), 31–40. https://doi.org/10.1145/3170358.3170421

[22] Paul Joseph-Richard and James Uhomoibhi. 2023. Which Data Sets Are Preferred by University Students in Learning Analytics Dashboards? A Situated Learning Theory Perspective. *INFORMS Transactions on Education* 24, 3 (2023), 220–237. https://doi.org/10.1287/ited.2023.0289

[23] Rogers Kaliisa, Kamila Misiejuk, Sonsoles López-Pernas, Mohammad Khalil, and Mohammed Saqr. 2024. Have Learning Analytics Dashboards Lived Up to the Hype? A Systematic Review of Impact on Students' Achievement, Motivation, Participation and Attitude. *ACM International Conference Proceeding Series* (2024), 295–304. https://doi.org/10.1145/3636555.3636884

[24] Fatemeh Salehian Kia, Stephanie D. Teasley, Marek Hatala, Stuart A. Karabenick, and Matthew Kay. 2020. How patterns of students dashboard use are related to their achievement and self-regulatory engagement. In *LAK20: 10th International Learning Analytics and Knowledge Conference.* 340–349. https://doi.org/10.1145/3375462.3375472

[25] Jeff J. Kosovich, Chris S. Hulleman, Kenneth E. Barron, and Steve Getty. 2015. A Practical Measure of Student Motivation: Establishing Validity Evidence for the Expectancy-Value-Cost Scale in Middle School. *Journal of Early Adolescence* 35, 5-6 (2015), 790–816. https://doi.org/10.1177/0272431614556890

[26] Alexandra Kuznetsova, Per B. Brockhoff, and Rune H.B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82, 13 (2017), 1–26. https://doi.org/10.18637/JSS.V082.I13

[27] Qiujie Li, Di Xu, Rachel Baker, Amanda Holton, and Mark Warschauer. 2022. Can student-facing analytics improve online students' effort and success by affecting how they explain the cause of past performance. *Computers & Education* 185, April (2022), 104517. https://doi.org/10.1016/j.compedu.2022.104517

[28] Lisa Lim, Srećko Joksimović, Shane Dawson, and Dragan Gašević. 2019. Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? *ACM International Conference Proceeding Series* (2019), 250–259. https://doi.org/10.1145/3303772.3303804

[29] Carol Midgley, Martin L Maehr, Ludmila Z Hruda, Eric Anderman, Lynley Anderman, Kimberley E Freeman, Margaret Gheen, Avi Kaplan, Revathy Kumar, Michael J Middleton, Jeanne Nelson, and Robert Roeser. 2000. *Manual for the Patterns of Adaptive Learning Sciences (PALS).* Technical Report. 74 pages.

[30] Fidelia A. Orji and Julita Vassileva. 2021. A Comparative Evaluation of the Effect of Social Comparison, Competition, and Social Learning in Persuasive Technology on Learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 12677 LNCS (2021), 369–375. https://doi.org/10.1007/978-3-030-80421-3{_}41

[31] Reinhard Pekrun. 2006. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review* 18, 4 (2006), 315–341. https://doi.org/10.1007/s10648-006-9029-9

[32] Reinhard Pekrun, Aisling Cusack, Kou Murayama, Andrew J. Elliot, and Kevin Thomas. 2014. The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction* 29 (2014), 115–124. https://doi.org/10.1016/j.learninstruc.2013.09.002

[33] Reinhard Pekrun, Andrew J. Elliot, and Markus A. Maier. 2009. Achievement Goals and Achievement Emotions: Testing a Model of Their Joint Relations With Academic Performance. *Journal of Educational Psychology* 101, 1 (2009), 115–135. https://doi.org/10.1037/a0013383

[34] Reinhard Pekrun, Herbert W. Marsh, Andrew J. Elliot, Kristina Stockinger, Raymond P. Perry, Elisabeth Vogl, Thomas Goetz, Wijnand A.P. van Tilburg, Oliver Lüdtke, and Walter P. Vispoel. 2023. A three-dimensional taxonomy of achievement emotions. *Journal of personality and social psychology* 124, 1 (2023), 145–178. https://doi.org/10.1037/pspp0000448

[35] Reinhard Pekrun, Raymond P Perry, Thomas Goetz, Kristina Stockinger, Herbert W Marsh, Raymond P Perry, and Herbert W Marsh. 2023. Achievement Emotions Questionnaire – Revised - User ' s Manual 2022 -. 124 (2023).

[36] Linda S. Perloff and Barbara K. Fetzer. 1986. Self-Other Judgments and Perceived Vulnerability to Victimization. *Journal of Personality and Social Psychology* 50, 3 (1986), 502–510. https://doi.org/10.1037/0022-3514.50.3.502

[37] P.R. R Pintrich, D. Smith, T. Garcia, and W. McKeachie. 1991. *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).* Technical Report.

[38] Emily Q. Rosenzweig, Allan Wigfield, and Jacquelynne S. Eccles. 2019. Expectancy-value theory and its relevance for student motivation and learning. *The Cambridge Handbook of Motivation and Learning* (2019), 617–644. https://doi.org/10.1017/9781316823279.026

[39] Peter J. Rousseeuw. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, C (1987), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[40] Holger Schielzeth and Shinichi Nakagawa. 2013. Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution* 4, 1 (2013), 14–24.

[41] Michael Schneider and Franzis Preckel. 2017. Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin* 143, 6 (2017), 565–600. https://doi.org/10.1037/bul0000098

[42] Corwin Senko, Chris S. Hulleman, and Judith M. Harackiewicz. 2011. Achievement Goal Theory at the Crossroads: Old Controversies, Current Challenges,

and New Directions. *Educational Psychologist* 46, 1 (2011), 26–47. https://doi.org/10.1080/00461520.2011.538646

[43] Sanam Shirazi Beheshitha, Marek Hatala, Dragan Gasevic, and Srecko Joksimovic. 2016. The Role of Achievement Goal Orientations When Studying Effect of Learning Analytics Visualizations. In *Proc. of the Sixth Int. Conference on Learning Analytics And Knowledge - LAK '16*. 54–63. https://doi.org/10.1145/2883851.2883904

[44] Teo Susnjak, Gomathy Suganya Ramaswami, and Anuradha Mathrani. 2022. Learning analytics dashboard: a tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education* 19, 1 (2022). https://doi.org/10.1186/s41239-021-00313-7

[45] Stephanie D. Teasley. 2017. Student Facing Dashboards: One Size Fits All? *Technology, Knowledge and Learning* 22, 3 (2017), 1–8. https://doi.org/10.1007/s10758-017-9314-3

[46] Heta Tuominen-Soini, Katariina Salmela-Aro, and Markku Niemivirta. 2011. Stability and change in achievement goal orientations: A person-centered approach. *Contemporary Educational Psychology* 36, 2 (2011), 82–100. https://doi.org/10.1016/j.cedpsych.2010.08.002

[47] Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and José Luis Santos. 2013. Learning analytics dashboard applications. *American Behavioral Scientist* 57, 10 (2013), 1500–1509.

[48] Tianyi Wang, Marek Hatala, Arash Ahrar, and Zhiquing Zheng. 2026. Linking Dashboard Elements and Retrospective Outcome Emotions in Dashboard Feedback : A Transmodal Analysis. In *Proceedings of The International Conference on Learning Analytics & Knowledge (LAK'26)*. 6 pages. https://doi.org/10.1145/3785022.3785090

[49] Joshua Weidlich, Aron Fink, Ioana Jivet, Jane Yau, Tornike Giorgashvili, Hendrik Drachsler, and Andreas Frey. 2024. Emotional and motivational effects of automated and personalized formative feedback: The role of reference frames. *Journal of Computer Assisted Learning* (2024), 1–18. https://doi.org/10.1111/jcal.13024

[50] Aaron C. Weidman, Jessica L. Tracy, and Andrew J. Elliot. 2016. The Benefits of Following Your Pride: Authentic Pride Promotes Achievement. *Journal of personality* 84, 5 (2016), 607–622. https://doi.org/10.1111/jopy.12184

[51] Bernard Weiner. 1985. An Attributional Theory of Achievement Motivation and Emotion. *Psychological Review* 92, 4 (1985), 548–573. https://doi.org/10.1037/0033-295X.92.4.548

[52] S. G. West, J. F. Finch, and P. J. Curran. 1995. Structural equation models with nonnormal variables. Problems and remedies. In *Structural equation modeling: Concepts, issues and applications*, R. H. Hoyle (Ed.). Sage, Newbury Park, CA, 56–75.

[53] Lisa A. Williams and David DeSteno. 2008. Pride and Perseverance: The Motivational Role of Pride. *Journal of Personality and Social Psychology* 94, 6 (2008), 1007–1017. https://doi.org/10.1037/0022-3514.94.6.1007

[54] Alyssa Friend Wise. 2014. Designing pedagogical interventions to support student use of learning analytics. *ACM International Conference Proceeding Series* (2014), 203–211. https://doi.org/10.1145/2567574.2567588

[55] Fabian Wolff, Friederike Helm, and Jens Möller. 2018. Testing the dimensional comparison theory: When do students prefer dimensional comparisons to social and temporal comparisons? *Social Psychology of Education* 21, 4 (2018), 875–895. https://doi.org/10.1007/s11218-018-9441-2

[56] Joanne V Wood. 1989. Theory and Research concerning social comparisons of personal attributes. *Psychological Bulletin* 106, 2 (1989), 231–248. https://doi.org/10.1037/0033-2909.106.2.231

[57] Joanne V. Wood. 1996. What is Social Comparison and How Should We Study it? *Personality and Social Psychology Bulletin* 22, 5 (1996), 520–537.

[58] Andrew Zamecnik, Vitomir Kovanović, Georg Grossmann, Srećko Joksimović, Gabrielle Jolliffe, David Gibson, and Abelardo Pardo. 2022. Team interactions with learning analytics dashboards. *Computers and Education* 185, April (2022). https://doi.org/10.1016/j.compedu.2022.104514

[59] Ethan Zell and Mark D. Alicke. 2010. The local dominance effect in self-evaluation: Evidence and explanations. *Personality and Social Psychology Review* 14, 4 (2010), 368–384. https://doi.org/10.1177/1088868310366144

[60] Ethan Zell and Jason E. Strickhouser. 2020. Comparisons Across Dimensions, People, and Time: On the Primacy of Social Comparison in Self-Evaluations. *Social Psychological and Personality Science* 11, 6 (2020), 791–800. https://doi.org/10.1177/1948550619884564

[61] Zhiquing Zheng, Marek Hatala, and Arash Ahrar. 2026. An Immediate And Retrospective Analysis Of Stability And Change: Examining Students' Emotion And Motivation Profiles In Response To Outcome-Presenting Learning Analytics Dashboards. In *Proceedings of The International Conference on Learning Analytics &Knowledge (LAK26)*. 10 pages. https://doi.org/10.1145/3785022.3785099