Logs or Self-Reports? Misalignment Between Behavioral Trace data and Surveys when Modeling Learner Achievement Goal Orientation

Logs or Self-Reports? Misalignment Between Behavioral Trace Data and Surveys When Modeling Learner Achievement Goal Orientation

Heeryung Choi
heeryung@mit.edu
Massachusetts Institute of Technology
Cambridge, MA, US

Philip H. Winne
winne@sfu.ca
Simon Fraser University
Burnaby, BC, Canada

Christopher Brooks
brooksch@umich.edu
University of Michigan
Ann Arbor, MI, US

Warren Li
liwarren@umich.edu
University of Michigan
Ann Arbor, MI, US

Kerby Shedden
kshedden@umich.edu
University of Michigan
Ann Arbor, MI, US

ABSTRACT

While learning analytics researchers have been diligently integrating trace log data into their studies, learners' achievement goals are still predominantly measured by self-reported surveys. This study investigated the properties of trace data and survey data as representations of achievement goals. Through the lens of goal complex theory, we generated achievement goal clusters using latent variable mixture modeling applied to each kind of data. Findings show significant misalignment between these two data sources. Self-reported goals stated before learning do not translate into goal-relevant behaviors tracked using trace data collected during learning activities. While learners generally articulate an orientation towards mastery learning in self-report surveys, behavioral trace data showed a higher incidence of less engaged learning activities. These findings call into question the utility of survey-based measures when up-to-date achievement goal data are needed. Our results advance methodological and theoretical understandings of achievement goals in the modern age of learning analytics.

CCS CONCEPTS

• **Applied computing** → **E-learning**; **Distance learning**.

KEYWORDS

achievement goals, trace data, survey data, latent variable mixture modeling

ACM Reference Format:
Heeryung Choi, Philip H. Winne, Christopher Brooks, Warren Li, and Kerby Shedden. 2022. Logs or Self-Reports? Misalignment Between Behavioral Trace Data and Surveys When Modeling Learner Achievement Goal Orientation. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3532106.3533449

LAK 2023, March 13–17, 2023, Arlington, TX, USA

ACM ISBN 978-1-4503-9358-4/22/06.
https://doi.org/10.1145/3532106.3533449

1 INTRODUCTION

Identifying learners' achievement goals [12, 15, 49] is often a first step in supporting their learning experiences. Successful self-regulated learning (SRL) requires continuous interactions with goals: selecting a goal (planning), operating to accomplish that goal (monitoring and regulating), and evaluating potential gaps between the goal and actual operations used to learn and outcomes achieved by those operations (reflection) [57, 62]. Understanding learners' goals is also important in providing feedback to help learners tune learning. Without understanding where learners are and aim to be, it is difficult to identify their learning trajectory and recommend a productive next step [22]. One major theory used to understand how learners navigate these issues is achievement goal theory [11, 13, 15].

Despite the key role of achievement goals in guiding learners' plans and actions, few studies have evaluated the limits of self-reported survey responses as representations of goals within dynamic learning contexts. Many studies have criticized limits of self-reported survey responses in capturing dynamic constructs relevant to SRL [2, 18, 52, 54, 61, 63]. The field of learning analytics is also increasingly aligning behavioral trace data with theoretical constructs as a complement to surveys. Behavioral trace data logged as clickstream interactions are designed to capture learner operations on information within educational software systems as learners learn. Trace data can model achievement goals in real time without interrupting learners with questions, thus closing potential gaps between self-reported goals before or after learning and a learner's dynamic expressions of achievement goal orientation during learning. Yet, achievement goals still are predominantly measured through self-reported surveys across domains [3, 6, 8, 17, 19, 26, 27, 39, 58].

In this study, we compared achievement goals self-reported in surveys before learning to traces unobtrusively gathered during learning. We used latent variable mixture modeling to identify achievement goal clusters represented in each data source. We found that self-reported goals did not directly manifest as online behavior. On prospective surveys, most learners identified their goals as mastery-oriented or a mixture of mastery- and performance-oriented. However, trace data characterized more than half of these learners as infrequently engaged with course materials designed to attract learners with those goals. This (1) raises significant questions about equating achievement goals self-reported before learning to achievement goals pursued during learning and (2) challenges the value of self-reports when up-to-date goal states are theorized to shape how learners learn. The value surveys may provide to understanding learners' expectations about how they plan to learn do not validly indicate adaptations to those expectations over time. This drawback could be important when instructors or researchers aim to support learners engaging with adaptive learning systems.

2 THEORETICAL FRAMING

2.1 Goal complex theory

While there are multiple versions of achievement goal theories, the theoretical base of this work is goal complex theory [12, 15, 49]. It proposed two dimensions of achievement goals: a *what* dimension and a *why* dimension. The *what* dimension is further divided into two categories: *mastery* and *performance*. Learners aiming for self-improvement which they judged by an intra-personal standard are considered *mastery* learners, while learners aiming to outperform others gauged by a normative standard such as a grade are deemed *performance* learners [13, 44].

Goal complex theory is particularly useful in clearly designing indicators of performance goals. Goal orientation theory, one of the earlier achievement goal theories, defined the motivation behind performance goals as a desire to demonstrate one's competence and gain recognition [1, 10, 35]. Yet, Hulleman et al. [25] showed that some studies either combined or replaced this original definition of performance goals with a more general motivation: to outperform others [11, 14, 15, 44]. The desire to demonstrate one's competence could be understood to reflect a social comparison component where learners want to stand out compared to others. However, a desire to outperform others does not inherently include that component. As long as one can attain competence to a satisfactory extent, that is enough. These two definitions of performance goals led researchers to debate which perspective should be dropped [7] and generated some confusion about the finding of using the same term 'performance goals' with different definitions.

Unlike goal orientation theory, the *why* dimension of the goal complex theory contrasts *autonomous* versus *controlling* motivations arising from self-determination theory [9, 40]. Goal orientation theory's avoidance-approach dimension established learners' motivation to approach a desired stimulus (approach) or to avoid an undesired stimulus (avoidance). Goal complex theory positioned motivation in a context of agency, reconceptualizing approach and avoidance as *controlling* motivation versus *autonomous* motivation. A controlling motivation is responsive to external pressures or tangible rewards. In learning environments, these learners are often keen to display their competence by outperforming others and gaining recognition for those accomplishments from peers, family, or instructors [9]. Autonomous motivations encourage learners to consider their accomplishment in alignment with personal values or satisfaction. Learners in this category, for example, may be interested in outperforming others to gain a feeling of achievement without a strong desire to demonstrate competence to others. This conceptualization of motivational orientation was supported by Urdan and Mestas using self-reports [49]. As the goal complex theory does not forward the approach-avoidance dimension, previous studies sometimes modified survey instruments originally designed according to goal orientation theory, and some studies have combined items reflecting approach and avoidance to measure each orientation or used only approach-based items [5, 38, 44, 50].

2.2 Survey and trace data

Several studies identified limitations of surveys when studying achievement goals [16, 47, 55]. One limitation is that survey data are not fine-grained enough to capture learners' temporally changing, context-specific goals. Several studies tried to narrow learning experience to specific contexts by adding qualifiers to survey items such as 'this semester' or 'this class' [16, 47]. Yet, these approaches still may be too coarse to capture varying contexts that could importantly influence learners' goals, such as the moment when a learner decides to focus on earning a passing grade rather than explore supplementary or advanced materials after the first quiz with an unexpectedly low grade.

Another inherent drawback of surveys is that they require learners' careful attention to and monitoring of recalled information to generate accurate descriptions of their goals or motivation. Winne [54] pointed out that researchers do not know how learners selectively sample experience forming a basis for what they report in surveys. Survey respondents apply a somewhat mysterious computational process to integrate multiple recalled experiences into one answer to a survey item. This process may be cognitively demanding and potentially biased for various reasons [21], including confirmation biases and concerns of social presentation (social desirability). Trace data, which do not require learners' constant attention, barely suffer these difficulties. Trace data can directly and automatically record learners' on-the-spot and dynamically adjusted behavior within and across specific contexts. The validity of interpretations of trace data depends on the qualities of a researcher's theory and attributes of the context in which a learner generates traces [28, 29, 56]. Rigorous attention to designing indicators and properties of methods for analyzing data can benefit studies using trace data.

Despite these limitations of survey data and perhaps due to challenges in coordinating survey data with trace data, research exploring instrumentation beyond surveys to measure achievement goals has been scant. Zhou and Winne [61] conducted one of the few studies using trace data to measure achievement goals alongside prospective surveys. They found weak to no correspondence across these two types of data despite both being designed to measure the same constructs. Zhou and Winne also found that trace data was a stronger predictor of participants' posttest achievement than were survey data. Considering that achievement goals have been predominantly studied using surveys, these are striking results.

While Zhou and Winne [61] shed some light on the validity issue relating to survey versus trace data as representations of achievement goals, their study also had limitations. First, the generalizability to authentic learning environments is likely limited. Their study was conducted in a lab where participants may not have a strong interest in or concern about learning. In such study settings, it can be questioned whether learners were genuinely motivated to develop new knowledge (mastery goals) or outperform other participants (performance goals). Zhou and Winne [61] also did not probe the causes of discrepancies between survey responses and trace data. The stronger correlation between trace data and posttest achievement, as theory predicts [54, 56], does not guarantee that trace data can be validly interpreted as measuring motivation. Learning analytics researchers should be cautious in advancing interpretations and forming analytics relating to learners' achievement goals. Finally, Zhou and Winne [61] did not investigate whether combining data sources might provide better ground for validly interpreting data intended to represent learners' achievement goals. Considering that each kind of instrumentation may have offsetting strengths and limitations, a natural step for the research is investigating whether combined data more strongly advance understanding of achievement goals and provides a sturdier ground for developing learning analytics.

3 STUDY OVERVIEW

Our work builds on Zhou and Winne's study [61] by conducting a field study collecting surveys and trace data from learners who engaged in authentic educational tasks associated with a credit-bearing course as part of an academic degree. We (1) examined information in the survey and trace data designed to capture achievement goals, (2) compared these two data sources to explore potential discrepancies between them, and (3) investigated whether combining these types of data was complementary to improving validity. Latent variable mixture modeling was applied to form clusters based on survey alone, trace data alone, and both types of data. We posed these research questions:

- RQ1. What goal clusters can be identified using survey data?
- RQ2. What goal clusters can be identified using trace data?
- RQ3. How different are goal clusters identified using combined surveys and trace data identified for RQ1 and RQ2?

Answering RQ1 and RQ2 contrasts patterns in learners' survey responses and behavioral trace data. This enriches understanding of how these data (1) differ in representing learners' goals and (2) align with the goal complex theory. RQ3 explores the possibility of a complementary combination of surveys and trace data in forming clusters.

3.1 Field Context

Survey and trace data were collected during two iterations of an introductory data science course offered in September 2021 (151 enrollments) and in January 2022 (98 enrollments). This course was the first programming course in an online Master's degree program in applied data science at the University of Michigan. The course had several characteristics distinguishing it from non-credentialed online courses and traditional residential college courses. The course was credit-bearing within a degree pathway where enrollment was limited and full tuition was required. The student-instructor ratio aligned to residential degree-granting programs (approximately 1 instructional aide per 50 enrolled students in addition to an instructor of record for the course), and individual synchronous office hours were readily available to students approximately daily. Compared to traditional residential college courses, learners in this course were more diverse in age, background knowledge, and level of education. Many were employed full-time, had parental responsibilities, and were part-time students in the university degree program.

In this four-week long, 1 credit unit course, learners had to submit four weekly mandatory assignments, each worth 25% of total course credit. Learners could earn additional credits toward their final grade by submitting bonus assignments. The top letter grade, *A+*, was awarded only to students submitting one or more bonus assignments and earning 100% in each mandatory assignment. The topic of the course, Data Manipulation, required that assignments be completed in the Python programming language. The curriculum covered introductory regular expressions, numerical python, and the `pandas` data manipulation toolkit. Learners could submit any particular assignment as many times as they wanted until the assignment deadline at the end of each week. Submissions were graded using an automatic code grading system. All course materials were released on the first day of the course, and learners could finish their course at their own pace as long as they submitted their weekly assignments by the end-of-week due date. The data collection timeline is shown in Figure 1.

4 METHODS

Survey data were collected to identify learners' expectations for achievement goals at two points in the course: the beginning of week 1 and the beginning of week 3. This allowed us to examine possible changes in expectations about achievement goals for the course's first half (weeks 1 and 2) and the second half (weeks 3 and 4) as learners gained experience with course content and pacing. Trace data were collected dynamically throughout the course. We separated trace data into one set collected in the first half of the course and another collected in the second half of the course.

4.1 Survey Instruments

Two surveys were used to collect self-reported achievement goals: the Achievement Goal Questionnaire (AGQ-R) [13] and a motivation survey from Vansteenkiste et al. [51]. Both approach and avoidance indicators were included and were not differentiated in the data analysis based on Vansteenkiste et al. [50]. Survey data included sixteen indicator variables: twelve items of the AGQ-R [13] plus four items from the questionnaire measuring motivations behind performance-oriented goals [46, 51]. Learners were informed their participation was optional and no reward was given for participation. Table 1 provides an overview of the survey sources, the relationship those sources had with theoretical constructs, and the specific variables we measured. The survey materials are available here[1]

4.1.1 Achievement Goal Questionnaire-Revised (AGQ-R). The Achievement Goal Questionnaire-Revised (AGQ-R) was designed by Elliot and Murayama [13] using American undergraduates' self-reported achievement goals about an examination in their college course. The questionnaire is composed of twelve 5-point Likert-scale items and these items were sorted into six mastery categorical indicators (three *mastery_approach* items and three *mastery_avoidance*

[1] https://osf.io/9n4ts

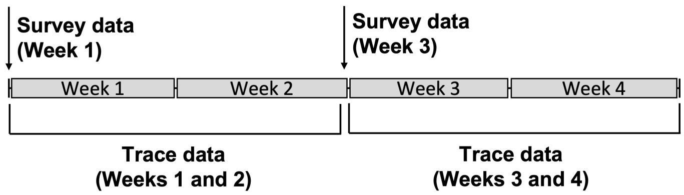

Figure 1: Course timeline for collecting survey data and trace data.

items) and six performance categorical indicators (three *performance_approach* items and three *performance_avoidance* items).

4.1.2 Motivation questionnaire. The motivation questionnaire was designed by Vansteenkiste et al. [51] based on Sheldon and Kasser [46]. It is composed of four 5-point Likert-scale items representing the external, introjected, identified, and intrinsic motivations behind performance goals. Based on a Vanssteenkiste et al.'s study [51], the external and introjected items were used as indicators for performance-controlling goals (two *performance_controlling* indicators); the identified and intrinsic items were used as indicators for performance-autonomous goals (two *performance_autonomous* indicators).

4.2 Trace Instruments

The first and the fourth authors collaborated with the course instructor to re-design several features of the curriculum to create indicators tailored to course objectives. To reduce the over- and under-representation of the achievement goal constructs for trace indicators, the domain modeling step of the Evidence-Centered Design (ECD) framework [4, 31, 32] was followed. ECD guides identifying alignment of indicators to targeted constructs, setting stronger foundations for validly interpreting finding of a study. In particular, the domain modeling step helps researchers represent how the design of a measure (1) obtains data-based evidence about the targeted knowledge and skills of learners, (2) supports claims based on that data-based evidence, and (3) considers possible counterclaims. Some of these are reported below, and the complete framework implementation results are available as supplementary material here[2]. In the next section the variable N in an indicator name represents the week in which data were collected to form the indicator. For example, *bonus1_sharing* indicates that indicators would exist for week 1 to measure learners' preference of sharing bonus assignments described in section 4.2.1. A full list of trace indicators created in this study is presented in Table 2.

4.2.1 Weekly bonus assignments. Optional weekly bonus assignments provided learners opportunities to earn additional credit. Learners were also asked to declare if and how they would like to share their work with faculty members and peers after the assignment deadline: not sharing, sharing anonymously, or sharing with their names. This indicator was designed to support performance-oriented learners who were interested in earning bonus credits and demonstrating performance to others. The first two sharing options were considered performance-autonomous goals. The third option was considered a performance-controlling goal since it sought recognition from peers and instructors more so than the learners choosing one of the first two options. Because these assignments were described as offering opportunities to practice skills already covered in the course rather than developing new skills, participating in bonus assignments was not considered an indicator of mastery goal orientation. From logs of these submissions, categorical indicator variables *bonusN_sharing* were formed with the following potential states: the bonus assignment not submitted (not relevant to any particular goal), bonus assignment submitted and not shared (performance-autonomous), bonus assignment submitted and anonymously shared (performance-autonomous), and bonus assignment submitted and shared with name declaration (performance-controlling).

4.2.2 Weekly extra assignment. Optional weekly extra assignments were designed to attract learners with a mastery orientation. Unlike bonus assignments, these assignments did not offer any additional credit and were designed to entice learners who were motivated to learn skills beyond the prescribed course curriculum. Binary indicators *extraN_submitted* indicated if learners submitted the extra assignment in the Nth week.

4.2.3 Biweekly tip-of-the-week email and Jupyter Notebook. In weeks 2 and 4, the course instructor released a tip-of-the-week Jupyter Notebook and sent a notification email to all students about the release. These assignments and emails targeted mastery-oriented learners. In each tip-of-the-week Jupyter Notebooks, the instructor explained how to write more efficient and readable Python code. Unlike weekly mandatory, bonus, and extra assignments, there were neither specific tasks assigned nor was a deadline set for this learning material. Every time a learner opened a tip-of-the-week Jupyter Notebook, added a cell, executed a cell successfully, executed a cell with error messages, removed a cell, and changed contents in a cell, an event log datum was generated with a timestamp. From these data, indicators *emailN_count* and *notebookN_count* were created. These were continuous variables counting the times learners opened the notification emails and interacted with each tip-of-the-week Jupyter Notebook.

4.2.4 Interactions with bonus and extra assignments. Learners might not have completed and submitted bonus or extra assignments due to time constraints or perceptions about the elevated difficulty of assignments despite a motivation to earn additional credits or learn advanced concepts. Some learners might simply have browsed these optional assignments out of curiosity without serious intention to complete them. To consider these different motivations behind an incomplete assignment, event log data were counted. The continuous indicators, named *bonusN_count* and *extraN_count*, indicated

[2]https://osf.io/cqxu6

Table 1: The survey indicators and their relationship to theoretical constructs.

Instrument	Theoretical Construct	Indicator Variable (n)
AGQ-R	Mastery	*mastery_approach* (3) *mastery_avoidance* (3)
	Performance	*performance_approach* (3) *performance_avoidance* (3)
Motivation questionnaire	Performance-controlling Performance-autonomous	*performance_controlling* (2) *performance_autonomous* (2)

performance and mastery goal pursuit, respectively. These were designed based on the rationale that, the stronger learners' motivation was to engage with assignments, the more active would be their engagement with the assignment.

4.2.5 Additional submissions of mandatory assignments. Some learners submitted their mandatory or optional assignments again even after they scored 100% on their assignment submission. While these additional submissions did not add any extra points to their final grade, some learners continued to experiment with alternative ways of coding. This behavioral pattern was used as another indicator of mastery goal orientation. The continuous indicators *additionalN_count* was the number of additional assignments learners submitted even after scoring 100%. For example, if a learner made additional submissions for the previously submitted week 1 mandatory assignment, the week 4 mandatory assignment, the week 1 bonus assignment, and the week 2 extra assignment, the indicator value for this learner was 4.

4.3 Data analysis

Two survey datasets were formed at the beginning of weeks 1 and 3 to investigate RQ1. For RQ2, two trace datasets were collected: the first one spanned weeks 1-2 of the course and the second covered weeks 3-4. For RQ3, two combined datasets were merged to form two datasets, one covering the first half of the course and the second covering the last half of the course. Figure 2 shows the flow of learner progress through the course and the relationship to the datasets.

We removed data for learners who did not submit all surveys, reducing sample size to 191 learners. First, confirmatory factor analysis was conducted to estimate a fit between each of the six cleaned datasets and goal complex theory. Then, latent variable mixture modeling [33] was applied on each dataset. In trace data, continuous variables were log-transformed to mitigate outliers and z-scored to improve the solution convergence process for estimating the parameters of each latent variable mixture modeling solution [24]. Mplus 8.6 [34] and the MplusAutomation R package [20] were used with a Maximum Likelihood with Robust standard errors (MLR) estimator. To answer RQ1, RQ2, and RQ3, each estimated cluster model was examined for statistical robustness and theoretical interpretability. The maximum number of clusters was limited to $1 \leq k \leq 6$ based on previous studies applying latent variable mixture modeling to survey responses describing achievement goals wherein none found more than six clusters [37, 42, 60]. In this analysis, each model M made up of k clusters is labeled as M_k. For example a model with 5 estimated clusters would be denoted as M_5.

Statistical criteria used to evaluate the model fit to each dataset were Log Likelihood (LL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), adjusted Bayesian Information Criterion (aBIC), Bootstrap Likelihood Ratio Test (BLRT) p-value, and Vuo-Lo-Mendell-Rubin likelihood ratio test (VLMR) p-value. These criteria are commonly used in studies applying latent variable mixture modeling [36, 37, 41, 53].

For each dataset, a best-fitting model was inspected to determine if the model resulted from normal termination of the estimation algorithm and that the best log-likelihood was replicated in runs from multiple independent starting values. This method ensured that the model could be deemed trustworthy instead of spurious [34]. Cluster sizes were also inspected by the authors to determine the quality of the final fitting model. While there is no agreement on appropriate cluster size limits, it has been noted that small clusters often do not add conceptual value or insight [53]. We followed suggestions from previous work [45, 53] and set a minimum bound of 5% (10 learners) in a cluster to ensure distinctive characteristics and increase generalizability. We planned to discard clusters smaller than this size, but none emerged from the analysis.

When warnings about a non-positive definite first-order derivative product matrix arose during estimation, which might imply less trustworthy model parameter estimates, indicators which caused that issue were pruned from the model. Often, a problematic indicator was a categorical variable representing a survey question which had the first (1-strongly disagree) or the last (5-strongly agree) Likert-scale option chosen by a very small number of learners. In this case, the response option was merged with the nearest neighbor (e.g., 1-strongly disagree was merged with 2-disagree), or the variable was discarded if merging options did not resolve the warning message. Then, we conducted latent variable mixture modeling with both the original model and the pruned model and investigated if the pruning process caused a major difference in analysis outcomes. Pruning indicators did not substantially alter the original model, thus the following discussion focuses on presenting the results of the pruned models for each dataset. Results of original models are available on the Open Science Foundation here[3].

5 RESULTS

Confirmatory factor analysis results showed that survey datasets had poorer fits with goal complex theory than trace datasets. Detailed results are here[4]. We focus on investigating information captured through each data source using latent variable mixture modeling.

[3]https://osf.io/a4wxr
[4]https://osf.io/cqxu6

Table 2: The trace indicators and their relationship to theoretical constructs.

Instrument	Theoretical Construct	Indicator Variable (N = Nth week)
Tip-of-the-week email	Mastery	*emailN_count*
Tip-of-the-week Jupyter Notebook	Mastery	*notebookN_count*
Bonus assignment	Performance-controlling	*bonusN_sharing* (sharing with names)
	Performance-autonomous	*bonusN_sharing* (not sharing, sharing anonymously)
	Performance	*bonusN_count*
Extra assignment	Mastery	*extraN_submitted*
	Mastery	*extraN_count*
Assignment	Mastery	*additionalN_count*

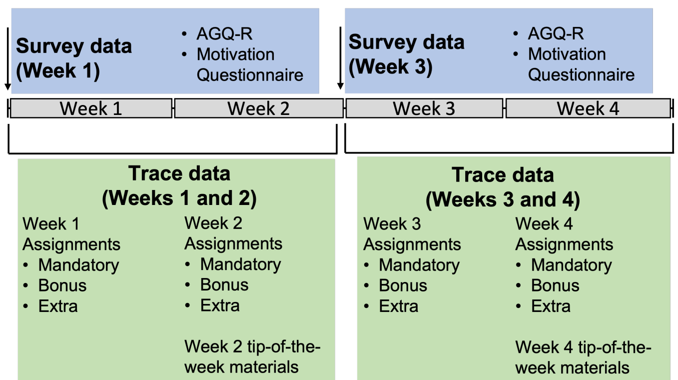

Figure 2: A course timeline showing which and when surveys and trace instruments were provided.

5.1 Model Fit

The best fitting models for the survey (RQ1) and trace (RQ2) indicators are shown in Table 3. The final fitting model M_3 for the combined dataset of the first half of the course did not report a conventionally statistically detectably better BLRT p-value and VLMR p-value ($p = 0.086$) than other solutions at $p = 0.05$ level and therefore was not further investigated.

5.2 Theoretical Interpretability: Surveys (RQ1)

Overall, survey-based solutions (RQ1) showed that learners generally agreed with most of the achievement goal items except performance-controlling items. This suggests that most learners identified them as mastery-oriented or both mastery- and performance-oriented. In the week 1 survey dataset, three clusters were identified (Figure 3). For the largest cluster ($n = 110$), we labeled it as 'positive' as learners in this cluster generally showed positive responses to all survey items except performance-controlling items. The second largest cluster ($n = 50$), which we labeled as 'strongly positive,' showed positive responses toward most of the survey items with the notable exception of the *performance_controlling1* indicator. The label chosen for the smallest cluster ($n = 31$) was 'mastery.' Learners in this cluster showed positive responses toward mastery items.

The week 3 survey dataset (Figure 4) showed a similar pattern. The same learners were clustered into two groups instead of three groups without a 'positive' group. We labeled the largest cluster ($n = 142$) 'mastery' in light of dominantly positive responses to mastery items. The other cluster ($n = 49$) was named as 'strongly positive.' These learners generally had positive responses to all items except performance-controlling ones.

5.3 Theoretical Interpretability: Traces (RQ2)

In contrast to the survey-based clusters showing high positivity for most achievement goals, trace-based solutions clustered many learners into 'less engaged' groups (RQ2). The best fitting model M_3 for the first half of the course from trace data (M_3) even showed that regardless of learners' cluster, learners barely engaged with the second-week bonus assignment (Figure 5). Due to the overall low engagement with the bonus assignment in the second week, we were unable to infer learners' motivations from sharing preference indicators. Learners in the largest cluster ($n = 95$) also rarely interacted with other optional course materials designed for mastery- and performance-oriented learners (Table 4). These learners also rarely made additional submissions once they reached scores of 100%. On the other hand, learners in the second largest cluster ($n = 52$), which we labeled 'mastery and performance', generated the highest counts of log data from engagement with week 1 bonus assignments (indicators *bonus1_count*), extra assignments (*extra1_count, extra2_count*), and tip-of-the-week Jupyter Notebook

Table 3: Final solutions per dataset.

Dataset (weeks 1-2)	Final model	Dataset (weeks 3-4)	Final model
Week 1 survey	M_3	Week 3 survey	M_2
Weeks 1 and 2 trace	M_3	Weeks 3 and 4 trace	M_3
Weeks 1 and 2 combination	-	Weeks 3 and 4 combination	M_2

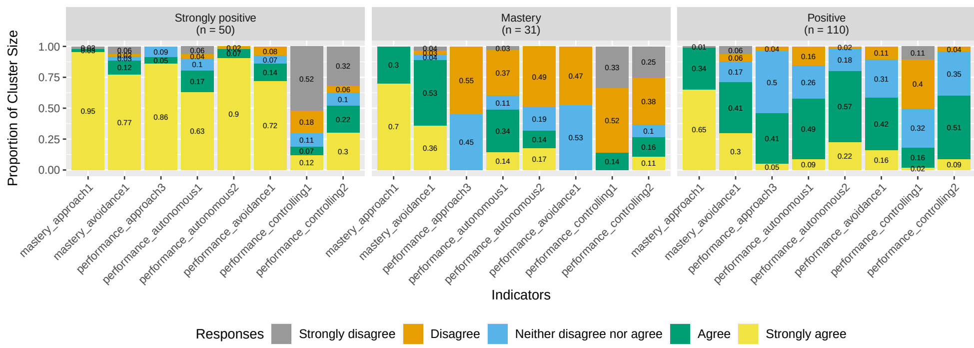

Figure 3: Response proportion (i.e., thresholds of categorical indicators) of the 3-cluster model on the week 1 survey dataset.

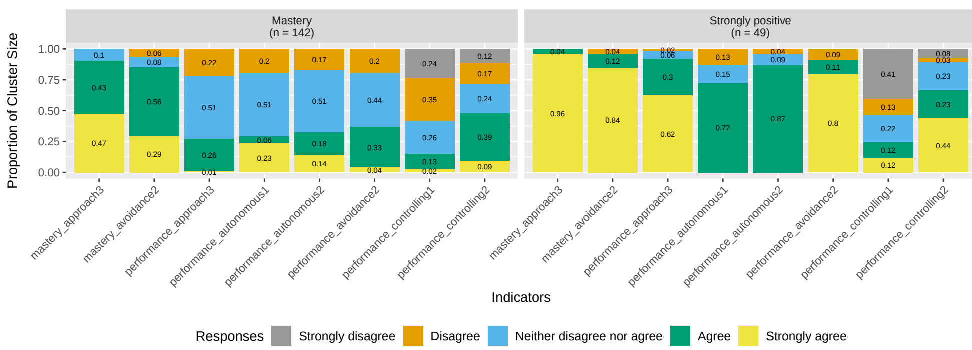

Figure 4: Response proportion (i.e., thresholds of categorical indicators) of the 2-cluster model on the week 3 survey dataset.

(*notebook2_count*). The smallest cluster was named ‘performance’ (n = 44). Learners in this cluster showed more targeted interest in materials bearing on their grades. While they opened tip-of-the-week notification emails more than other learners (*email2_count*), their engagement with tip-of-the-week Jupyter Notebook (*notebook2_count*) was lower than others. Their engagement with extra assignments (*extra1_count*, *extra2_count*) was also much lower than the engagement with bonus assignments (*bonus1_count*, *bonus2_count*), the latter of which had the potential to increase grades.

M_3 for the weeks 3-4 trace dataset showed that the largest cluster (n = 135) was also composed of less engaged learners. The second largest cluster was named ‘performance and weak mastery’ (n = 43). Learners in this cluster selectively engaged with the third-week bonus assignments (*bonus3_count*) while showing low engagement with extra assignments in the same week (*extra3_count*). They also showed weak mastery behaviors through slightly higher engagement with tip-of-the-week notebooks (*notebook4_count*) and

Table 4: Means (standard deviations) of continuous variables in weeks 1-2 trace dataset.

	Less engaged (n = 95)	Mastery and performance (n = 52)	Performance (n = 44)
email2_count	-0.058 (0.094)	-0.013 (0.139)	0.185 (0.180)
notebook2_count	-0.280 (0.074)	0.587 (0.168)	0.145 (0.156)
bonus1_count	0.987 (0.003)	1.022 (0.033)	0.937 (0.042)
bonus2_count	0.443 (0.083)	0.749 (0.115)	0.343 (0.153)
extra1_count	-0.641 (0.012)	1.692 (0.057)	-0.393 (0.008)
extra2_count	-0.296 (0.051)	0.792 (0.206)	-0.087 (0.130)
additional12_count	-0.185 (0.064)	0.137 (0.163)	0.144 (0.179)

Note. Means were computed with z-scored continuous variables.

counts of additional submissions (*additional34_count*) than learners in the ‘less engagement’ group. Learners in this group did not

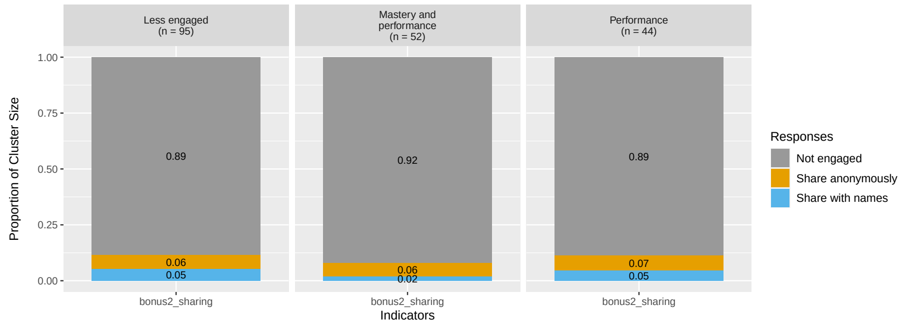

Figure 5: Proportions of engagement with learning materials (i.e., thresholds of categorical indicators) for the 3-cluster model on the weeks 1-2 trace dataset.

show strong preferences between anonymous sharing and sharing with names (*bonus4_sharing*). On the other hand, learners in the smallest 'mastery and performance-controlling' cluster ($n = 13$) highly interacted not only with bonus assignments but also with extra assignments and tip-of-the-week notebooks. They also made many additional submissions even after scoring 100% on mandatory assignments. Finally, they showed a slightly higher preference for sharing assignments with names.

5.4 Combined Data (RQ3)

Given that the latent variable mixture modeling failed to find the best fit model from the combination of data in the first half of the course, only the combined model for the second half of the course was considered (RQ3). The first three columns of Table 6 show the summary of estimated clusters in the final fitting model. The trace cluster output identifying three clusters with 135, 43, and 13 learners each was highly similar to the fitted model of the combined dataset, which had two clusters with 134 and 57 learners each. Cramer's V which measures the similarity between categorical variables was 0.987, which is close to the maximum value of 1 representing the complete association. Naturally, it did not show any detectable improvement from the model fit using trace data alone. On the other hand, Cramer's V showed a low association between survey data and the combined data.

6 DISCUSSION AND FUTURE WORK

In the context of an authentic, credit-bearing course, learners responding to survey items sorted into clusters reflecting predominantly mastery-oriented or mastery- plus performance-oriented motives, all of whom expressed dominantly positive responses. In contrast, trace-based clusters indicated that more than half of learners in the early and latter halves of the course clustered into a less-engaged group. The proportion of these less engaged learners sharply increased from the first half of the course (trace dataset 1) to the second half (trace dataset 2). That is, learners' self-reports did not translate into behaviors in this field study. This discrepancy raises significant issues for learning analytics researchers who consider learners' motivation an important predictor of achievement or an outcome in its own right.

One possible explanation for this discrepancy, at least at the beginning of the course, may be a difference between learners' expectations and direct experiences. As is common in many studies, we collected survey responses before the course began. These responses best reflect learners' expectations for achievement goals based on weakly grounded assumptions about the course and perhaps socially desirable response bias. It is reasonable that learners would have little basis for accurately predicting actual challenges and opportunities in the course. Revising goal orientations is both likely and rational. However, this shift suggests learning analytics using pre-course perceptions may rest on weak grounds for predicting early engagement with the course and performance during that period.

At the midway point of the course, this discrepancy between the third-week survey model and the second-half trace model was not reduced. Many learners continued to self-report combined mastery- and performance-oriented goals or predominantly mastery-oriented goals. In contrast, trace data gathered during the second half of the course indicated that more than half of learners did not engage with learning opportunities representing a theory-based description of mastery-oriented learners. Learners may have recognized that they had not met their initial goals and decided to try again. If that was the case, their attempts were not successful given what the trace data indicated. Or, perhaps learners did not seriously commit to goals that they reported to surveys. In either case, considering the increase of less engaged learners in the second half of the course, a goal "in mind" differed from a goal expressed "in action." Finally, survey instruments may elicit aspirational goals instead of realistic goals. In this case, learners responded to the survey describing the "person they would like to be" instead of based on the "person they will be." While aspirations perhaps should not be shattered, instructors nonetheless need to attend to learners' actual learning activities since intentions to learn can be realized only when productive activities are engaged. Again, learning analytics based on self-report data would potentially misrepresent learning activities and poorly predict learning outcomes arising from how learners actually engage in their coursework.

This discrepancy invites questions about common interpretations of achievement goal theory and commonly used self-report instruments as indicators of motivation. Multiple studies reported that learners favor mastery goals over performance goals when

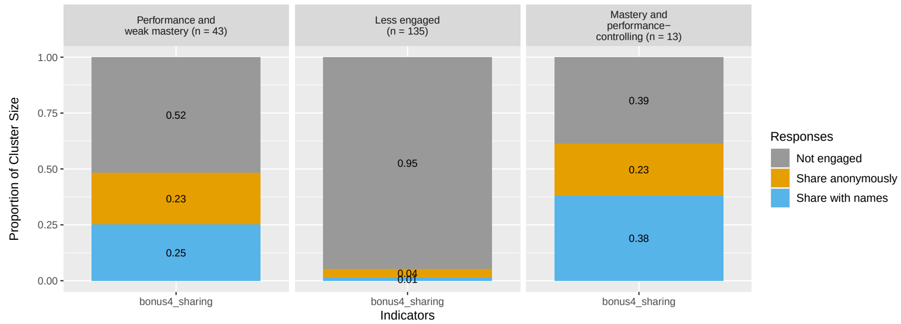

Figure 6: Proportions of engagement with learning materials (i.e., thresholds of binary indicators) of the 3-cluster model on the weeks 3-4 trace dataset.

Table 5: Means (standard deviations) of continuous variables in weeks 3-4 trace dataset.

	Performance and weak mastery (n = 43)	Less engaged (n = 135)	Mastery and performance-controlling (n = 13)
email4_count	-0.031 (0.139)	0.009 (0.087)	0.313 (0.377)
notebook4_count	0.382 (0.205)	-0.149 (0.066)	0.386 (0.397)
bonus3_count	1.537 (0.071)	-0.602 (0.020)	1.326 (0.123)
extra3_count	-0.203 (0.047)	-0.279 (0.006)	3.562 (0.245)
additional34_count	0.246 (0.190)	-0.197 (0.057)	0.885 (0.461)

Note. Means were computed with z-scored continuous variables.

Table 6: Cramer's V correlation of combination cluster outputs with (1) survey and (2) trace.

Dataset	Survey cluster output	Trace cluster output	Combination cluster output	Cramer's V (with survey)	Cramer's V (with trace)
Week 3 and 4	142-49	135-43-13	134-57	0.062	0.987

responding in interviews or to surveys e.g., [7, 23, 48]. It has been assumed learners' achievement goals measured before learning were maintained in subsequent learning activities, laying groundwork to infer causal relationships between achievement goals and learning outcomes [26, 30, 43, 59]. Beyond the famous maxim that "correlation is not causation" and the absence of a consistent correlation between achievement goals and academic achievement for mastery-focused learners [43], our study suggests many learners stating mastery-oriented goals based on surveys before learning may be misleading if this is taken to indicate how learners will actually go about learning, the proximal causal factors that lead to the achievement of lack thereof. This is a critical issue, in particular, for an adaptive learning system which depends on up-to-date learner data.

The strong similarity of clusters based on the combined data to trace-based clusters implies that survey data neither contributed to the cluster outputs nor contradicted trace data. This supports interpreting that self-report surveys may not add to identifying learners' goals; there was no complementary effect in combining the two data source. It also raises questions about roles for self-report data beyond representing learners' expectations about motivation.

One fundamental implication for future research is to clearly and operationally define what 'an achievement goal' is. Is it: learners' (1) *expectations* about learning, (2) *enactments* in accord with stated goals, and/or (3) *memories* about goals and enactments after learning. The first and the third approaches to operationally defining goals will be valuable for researchers to understand learners' *perceptions* of their achievement goals as represented by self-report instruments: prospective and retrospective surveys, and think-aloud protocols. However, if researchers aim to identify the goal-relevant *behaviors* during learning, traces may be more useful than self-report instruments in validly inferring learners' motivation-in-action.

An implication for instructors is that such discrepancies between trace and survey data could be useful for identifying learners potentially needing support for self-regulated learning. For example, instructors observing particular profiles in self-report data might want to intervene in learners to encourage different forms of engagement with course materials than what self-reports might suggest would be the case. Our data also modestly suggest that learners might have made a goal adjustment as learners get more familiar with course features such as task difficulty. Much research remains to be done to more validly guide instructors' choices about how to support learners in such situations.

We note several limitations in our designs of behavioral trace indicators. The *bonusN_sharing* indicators might not have fully distinguished learners holding different performance motivations since some controlling learners could have chosen to share answers anonymously instead of revealing their names if they did not feel confident about their work and wanted to avoid displaying that to peers. This study also did not consider traces beyond instrumentation within the course. It is possible that learners in the 'less engaged' cluster pursued goals in ways our data could not detect. For example, learners might have read blog articles or followed tutorials outside of the course instead of engaging with additional materials we provided and experimenting with alternative answers even after receiving 100% on an assignment. A future study with a broader range of instrumentation could explore how online learners seek to accomplish goals outside the boundaries of their courses.

7 CONCLUSION

Through latent variable modeling, we observed a notable discrepancy between learners' self-reported goals before learning and actual behaviors for pursuing goals during learning, even after learners had experience with half of the course. Learners mostly identified themselves as mastery learners on surveys whereas more than half of them did not demonstrate mastery-oriented behaviors and showed low engagement. Even among learners who actively pursued goals, many were performance-oriented learners. Furthermore, when trace data and surveys were combined, survey data made practically no contribution to clustering results. These findings raise questions on the widespread practices in the learning analytics community regarding categorizing learners' goals based on prospective survey responses and trying to interpret correlations between prospective self-reported goal orientations and learning outcomes. This study prompts future research to give sharper attention to methodological and theoretical accounts of achievement goals. In particular, the study raises issues relating to data for developing learning analytics that helps learners and their instructors (1) identify goals and (2) coordinate goals with learning activities to increase learning success.

REFERENCES

[1] Carole Ames. 1992. Classrooms: Goals, structures, and student motivation. *J. Educ. Psychol.* 84, 3 (1992), 261.

[2] Roger Azevedo. 2015. Defining and Measuring Engagement and Learning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues. *Educ. Psychol.* 50, 1 (Jan. 2015), 84–94.

[3] James W Beck and Aaron M Schmidt. 2013. State-level goal orientations as mediators of the relationship between time pressure and performance: a longitudinal study. *J. Appl. Psychol.* 98, 2 (March 2013), 354–363.

[4] John T Behrens, Philip Piety, Kristen E DiCerbo, and Robert J Mislevy. 2018. Inferential foundations for learning analytics in the digital ocean. *Learning analytics in education* (2018), 1–48.

[5] Moti Benita, Lennia Matos, and Yasmin Cerna. 2021. The effect of mastery goal-complexes on mathematics grades and engagement: The case of Low-SES Peruvian students. *Learning and Instruction* (Oct. 2021), 101558.

[6] George Botsas and Susana Padeliadu. 2003. Goal orientation and reading comprehension strategy use among students with and without reading difficulties. *Int. J. Educ. Res.* 39, 4 (Jan. 2003), 477–495.

[7] Jere Brophy. 2005. Goal Theorists Should Move on From Performance Goals. *Educ. Psychol.* 40, 3 (Sept. 2005), 167–176.

[8] Zhi-Hong Chen, Calvin C Y Liao, Hercy N H Cheng, Charles Y C Yeh, and Tak-Wai Chan. 2012. Influence of game quests on pupils' enjoyment and goal-pursuing in math learning. *Journal of Educational Technology & Society* 15, 2 (2012), 317–327.

[9] Edward L Deci and Richard M Ryan. 2000. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychol. Inq.* 11, 4 (2000), 227–268.

[10] Carol S Dweck. 1986. Motivational processes affecting learning. *Am. Psychol.* 41, 10 (1986), 1040.

[11] Andrew J Elliot. 1999. Approach and avoidance motivation and achievement goals. *Educ. Psychol.* 34, 3 (June 1999), 169–189.

[12] Andrew J Elliot. 2005. A conceptual history of the achievement goal construct. In *Handbook of competence and motivation*, Andrew J Elliot and Carol S Dweck (Eds.). Vol. 16. New York: The Guilford Press, 52–72.

[13] Andrew J Elliot and Kou Murayama. 2008. On the measurement of achievement goals: Critique, illustration, and application. *J. Educ. Psychol.* 100, 3 (2008), 613.

[14] Andrew J Elliot, Kou Murayama, and Reinhard Pekrun. 2011. A 3 × 2 achievement goal model. *J. Educ. Psychol.* 103, 3 (2011), 632–648.

[15] Andrew J Elliot and Todd M Thrash. 2001. Achievement goals and the hierarchical model of achievement motivation. *Educ. Psychol. Rev.* 13, 2 (2001), 139–156.

[16] Sara J Finney, Suzanne L Pieper, and Kenneth E Barron. 2004. Examining the Psychometric Properties of the Achievement Goal Questionnaire in a General Academic Context. *Educ. Psychol. Meas.* 64, 2 (April 2004), 365–382.

[17] Joanna Giota and Daniel Bergh. 2021. Adolescent Academic, Social and Future Achievement Goal Orientations: Implications for Achievement by Gender and Parental Education. *Scandinavian Journal of Educational Research* 65, 5 (July 2021), 831–850.

[18] Jeffrey A Greene and Roger Azevedo. 2010. The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educ. Psychol.* 45, 4 (2010), 203–209.

[19] Leslie Morrison Gutman. 2006. How student and parent goal orientations and classroom goal structures influence the math achievement of African Americans during the high school transition. *Contemp. Educ. Psychol.* 31, 1 (Jan. 2006), 44–63.

[20] Michael N Hallquist and Joshua F Wiley. 2018. MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Struct. Equ. Modeling* 25, 4 (Jan. 2018), 621–638.

[21] R Hastie. 2001. Problems for judgment and decision making. *Annu. Rev. Psychol.* 52 (2001), 653–683.

[22] John Hattie and Helen Timperley. 2007. The Power of Feedback. *Rev. Educ. Res.* 77, 1 (March 2007), 81–112.

[23] Daphne Hijzen, Monique Boekaerts, and Paul Vedder. 2007. Exploring the links between students' engagement in cooperative learning, their goal preferences and appraisals of instructional conditions in the classroom. *Learning and Instruction* 17, 6 (Dec. 2007), 673–687.

[24] T George Hornby, Christopher E Henderson, Carey L Holleran, Linda Lovell, Elliot J Roth, and Jeong Hoon Jang. 2020. Stepwise Regression and Latent Profile Analyses of Locomotor Outcomes Poststroke. *Stroke* 51, 10 (Oct. 2020), 3074–3082.

[25] Chris S Hulleman, Sheree M Schrager, Shawn M Bodmann, and Judith M Harackiewicz. 2010. A meta-analytic review of achievement goal measures: different labels for the same constructs or different constructs with similar labels? *Psychol. Bull.* 136, 3 (May 2010), 422–449.

[26] Leong Yeok Jang and Woon Chia Liu. 2012. 2 × 2 Achievement goals and achievement emotions: a cluster analysis of students' motivation. *European Journal of Psychology of Education* 27, 1 (March 2012), 59–76.

[27] Stefan Janke and Oliver Dickhäuser. 2019. A neglected tenet of achievement goal theory: Associations between life aspirations and achievement goal orientations. *Pers. Individ. Dif.* 142 (May 2019), 90–99.

[28] Il-Hyun Jo, Dongho Kim, and Meehyun Yoon. 2014. Analyzing the Log Patterns of Adult Learners in LMS Using Learning Analytics. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* (Indianapolis, Indiana, USA) *(LAK '14)*. ACM, New York, NY, USA, 183–187.

[29] Vitomir Kovanovic, Dragan Gašević, Shane Dawson, Srećko Joksimovic, Ryan S Baker, and Marek Hatala. 2015. Does Time-on-task Estimation Matter? Implications on Validity of Learning Analytics Findings. *1* 2, 3 (2015), 81–110.

[30] Wenshu Luo, Scott G Paris, David Hogan, and Zhiqiang Luo. 2011. Do performance goals promote learning? A pattern analysis of Singapore students' achievement goals. *Contemp. Educ. Psychol.* 36, 2 (April 2011), 165–176.

[31] Robert J Mislevy and Geneva D Haertel. 2007. Implications of evidence-centered design for educational testing. *Educ. Meas. Issu. Pr.* 25, 4 (April 2007), 6–20.

[32] R J Mislevy, L S Steinberg, and others. 2003. Focus article: On the structure of educational assessments. *research and perspectives* (2003).

[33] Bengt O Muthén. 2001. Latent variable mixture modeling. In *New developments and techniques in structural equation modeling*. Psychology Press, 21–54.

[34] Bengt O Muthén and Linda Muthén. 2017. *Mplus*. Chapman and Hall/CRC.

[35] John G Nicholls. 1984. Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychol. Rev.* 91, 3 (July 1984), 328–346.

[36] Karen Nylund-Gibson and Andrew Young Choi. 2018. Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science* 4, 4 (Dec. 2018), 440–461.

[37] Dena A Pastor, Kenneth E Barron, B J Miller, and Susan L Davis. 2007. A latent profile analysis of college students' achievement goal orientation. *Contemp. Educ. Psychol.* 32, 1 (Jan. 2007), 8–47.

[38] Caroline Julia Pulfrey, Maarten Vansteenkiste, and Aikaterina Michou. 2019. Under Pressure to Achieve? The Impact of Type and Style of Task Instructions on Student Cheating. *Front. Psychol.* 10 (Aug. 2019), 1624.

[39] Richard Remedios and John T E Richardson. 2013. Achievement goals in adult learners: evidence from distance education. *Br. J. Educ. Psychol.* 83, Pt 4 (Dec. 2013), 664–685.

[40] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 1 (Jan. 2000), 68–78.

[41] James B Schreiber. 2017. Latent Class Analysis: An example for reporting results. *Res. Social Adm. Pharm.* 13, 6 (Nov. 2017), 1196–1201.

[42] Malte Schwinger and Elke Wild. 2012. Prevalence, stability, and functionality of achievement goal profiles in mathematics from third to seventh grade. *Contemp. Educ. Psychol.* (2012).

[43] Corwin Senko, Chris S Hulleman, and Judith M Harackiewicz. 2011. Achievement Goal Theory at the Crossroads: Old Controversies, Current Challenges, and New Directions. *Educ. Psychol.* 46, 1 (Jan. 2011), 26–47.

[44] Corwin Senko and Katie L Tropiano. 2016. Comparing three models of achievement goals: Goal orientations, goal standards, and goal complexes. *J. Educ. Psychol.* 108, 8 (Nov. 2016), 1178–1192.

[45] Lilly Shanahan, William E Copeland, Carol M Worthman, Alaattin Erkanli, Adrian Angold, and E Jane Costello. 2013. Sex-differentiated changes in C-reactive protein from ages 9 to 21: the contributions of BMI and physical/sexual maturation. *Psychoneuroendocrinology* 38, 10 (2013), 2209–2217.

[46] Kennon M Sheldon and Tim Kasser. 1995. Coherence and congruence: two aspects of personality integration. *J. Pers. Soc. Psychol.* 68, 3 (March 1995), 531–543.

[47] Kamden K Strunk. 2014. A Factor Analytic Examination of the Achievement Goal Questionnaire–Revised Supports a Three-Factor Model. *Psychol. Rep.* 115, 2 (Oct. 2014), 400–414.

[48] Tim Urdan. 2004. Using Multiple Methods to Assess Students' Perceptions of Classroom Goal Structures. *Eur. Psychol.* 9, 4 (Jan. 2004), 222–231.

[49] Tim Urdan and Miranda Mestas. 2006. The goals behind performance goals. *J. Educ. Psychol.* 98, 2 (May 2006), 354–365.

[50] Maarten Vansteenkiste, Willy Lens, Andrew J Elliot, Bart Soenens, and Athanasios Mouratidis. 2014. Moving the Achievement Goal Approach One Step Forward: Toward a Systematic Examination of the Autonomous and Controlled Reasons Underlying Achievement Goals. *Educ. Psychol.* 49, 3 (July 2014), 153–174.

[51] Maarten Vansteenkiste, Stijn Smeets, Bart Soenens, Willy Lens, Lennia Matos, and Edward L Deci. 2010. Autonomous and controlled regulation of performance-approach goals: Their relations to perfectionism and educational outcomes. *Motiv. Emot.* 34, 4 (Dec. 2010), 333–353.

[52] Marcel V J Veenman and Dorit van Cleef. 2018. Measuring metacognitive skills for mathematics: students' self-reports versus on-line assessment methods. *ZDM* (2018), 1–11.

[53] Bridget E Weller, Natasha K Bowen, and Sarah J Faubert. 2020. Latent Class Analysis: A Guide to Best Practice. *J. Black Psychol.* 46, 4 (May 2020), 287–311.

[54] Philip H Winne. 2010. Improving Measurements of Self-Regulated Learning. *Educ. Psychol.* 45, 4 (Oct. 2010), 267–276.

[55] Philip H Winne. 2020. Commentary: A Proposed Remedy for Grievances about Self-Report Methodologies. *Frontline Learning Research* 8, 3 (2020), 164–173.

[56] Philip H Winne. 2020. Construct and consequential validity for learning analytics based on trace data. *Comput. Human Behav.* 112 (Nov. 2020), 106457.

[57] Philip H Winne and Allyson F Hadwin. 1998. Studying as self-regulated learning. *Metacognition in educational theory and practice* 93 (1998), 27–30.

[58] Christopher A Wolters. 2004. Advancing Achievement Goal Theory: Using Goal Structures and Goal Orientations to Predict Students' Motivation, Cognition, and Achievement. *J. Educ. Psychol.* 96, 2 (June 2004), 236–250.

[59] Stephanie Virgine Wormington and Lisa Linnenbrink-Garcia. 2016. A New Look at Multiple Goal Pursuit: the Promise of a Person-Centered Approach. *Educ. Psychol. Rev.* 29, 3 (Feb. 2016), 407–445.

[60] Ying Zhang, Rainer Watermann, and Annabell Daniel. 2016. Are multiple goals in elementary students beneficial for their school achievement? A latent class analysis. *Learn. Individ. Differ.* 51 (Oct. 2016), 100–110.

[61] Mingming Zhou and Philip H Winne. 2012. Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction* 22, 6 (Dec. 2012), 413–419.

[62] Barry J Zimmerman. 2000. Self-Efficacy: An Essential Motive to Learn. *Contemp. Educ. Psychol.* 25, 1 (Jan. 2000), 82–91.

[63] Barry J Zimmerman. 2008. Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *Am. Educ. Res. J.* 45, 1 (March 2008), 166–183.