Meta-Psychology, 2024, vol 8, MP.2021.2918
https://doi.org/10.15626/MP.2021.2918
Article type: Replication Report

Open data: Yes
Open materials: Yes
Open and reproducible analysis: Yes
Open reviews and editorial process: Yes
Preregistration: Yes

Edited by: Thomas Nordström
Reviewed by: Stefan Gustafson, Kazuhisa Takemura
Analysis reproduced by: Lucija Batinović
All supplementary files can be accessed at OSF:
https://doi.org/10.17605/OSF.IO/VKSN2

Associations between Goal Orientation and Self-Regulated Learning Strategies are Stable across Course Types, Underrepresented Minority Status, and Gender: A Replication of Muis and Franco (2009)

Brendan A. Schuetze[1] and Veronica X. Yan[1]
[1]Educational Psychology Department, The University of Texas at Austin

In this pre-registered replication of findings from Muis and Franco [2009; Contemporary Educational Psychology, 34(4), 306-318], college students (N = 978) from across the United States and Canada were surveyed regarding their goal orientations and learning strategies. A structural equation modelling approach was used to assess the associations between goal orientations and learning strategies. Six of the eight significant associations (75%) found by Muis and Franco replicated successfully in the current study. Mastery approach goals positively predicted endorsement of all learning strategies (Rehearsal, Critical Thinking, Metacognitive Self-Regulation and Elaboration). Performance avoidance goals negatively predicted critical thinking, while positively predicting metacognitive self-regulation and rehearsal. Evidence for moderation by assignment type was found. No evidence of the moderation of these associations by gender, underrepresented minority status, or course type (STEM, Humanities, or Social Sciences) was found. The reliability of common scales used in educational research and issues concerning the replication of studies using structural equation modeling are discussed.

Keywords: SEM Replication; Achievement Goals; Learning Strategies; Heterogeneity

Motivational beliefs and learning strategies are a key theoretical mechanism of academic achievement, with interest to learning scientists and educators alike. In recent years, research in both of these areas has increasingly grappled with findings that motivational variables (Fong et al., 2019; Meece et al., 2006) and learning strategies (Deemer, 2004) often behave differently across socio-cultural contexts and identities. In this study, we present a pre-registered replication of Muis and Franco's (2009) study of the associations between goal orientation theory and learning strategies, with an extended set of analyses focused on determining the extent to which these associations are moderated by contextual (course type, assignments) and demographic (gender, underrepresented minority status) variables.

Goal Orientation Theory

Goal orientation theory has been a productive way of understanding academic motivation by examining the type of goals that students adopt as a result of their personal background and classroom context (Meece et al., 2006). In its most common formulation, goal orientation theory highlights that students can be oriented toward two types of goals (mastery or performance) an that the goals can be positively or negatively valenced (approach or avoidance), thus creating four general goal orientations: performance-approach, mastery-approach, performance-avoidance, and mastery-avoidance (see Elliot and McGregor, 2001).

Performance goals center around an individual's performance in terms of grades and relative standing among their peers. Mastery goals revolve around the state of one's knowledge. Approach orientation refers to the pursuit of positive outcomes (e.g., a high grade or deep understanding), while avoidance orientation means seeking to avoid negative outcomes (e.g., failing a class, misunderstanding a lecture or forgetting material). Putting these two axes (performance-mastery and approach-avoid) together, the meaning of the four broad goal orientations is clear. For example, performance-approach refers to the aim of achieving better performance than fellow students, while mastery-avoidance refers to the goal of avoiding having not mastered the course content.

A multitude of studies have found that higher lev-

els of self-reported approach-goal orientations correlate with (Senko et al., 2011), and also—in the case of intervention studies (e.g., Fuchs et al., 1997)—cause (Kaplan and Maehr, 2007) higher student performance. Ultimately, it seems that approach goals are positively associated with performance, regardless of whether they are performance or mastery in orientation and holding both types of goals simultaneously can contribute to academic achievement (Harackiewicz et al., 1997; Senko et al., 2013).

Is Time-on-Task the Mechanism?

Although the broader association between approach goals and academic achievement appear relatively robust across studies (Harackiewicz et al., 1997; Senko et al., 2013), less is known about the mechanism by which motivational states cause higher performance (as also discussed in the growth mindset literature, Yan and Schuetze, 2023). Some researchers, such as Tabak et al. (2009), have put forth theoretical models showing that motivational beliefs work on academic performance primarily through the route of increased time-on-task (also called "task engagement," "persistence," "effort regulation" and "effort management"; Liem et al., 2008). This hypothesis seems plausible. Since the early work of J.B. Carroll (Carroll, 1963, 1989), time-on-task has been known to be a key predictor of achievement in both academic (e.g., Lee, 2018) and non-academic (Macnamara et al., 2014) domains. It does not seem intuitively necessary that higher achievement motivation causes students to act in a qualitatively different manner. Rather, motivated students may simply spend more time than students with lower achievement motivation, but in a qualitatively similar manner (Yan and Schuetze, 2023).

Quality, not just Quantity of Study

However, other researchers, particularly those who study motivation using a goal orientation framework, have argued that motivated students will use more beneficial strategies (Engelschalk et al., 2017; Senko et al., 2011). In other words, increased motivation benefits the quality and quantity of study. Under this hypothesis, students with higher approach (both mastery and performance) goal endorsement will preferentially engage in deeper learning strategies, which are also linked with higher achievement (Muis and Franco, 2009). Deeper learning strategies are those that engage students in critical thinking, monitoring comprehension, organizing information, and drawing connections between educational materials. These strategies are thought to be beneficial for long-term learning, as opposed to shallow strategies, such as rehearsal/repetition, underlining, or highlighting (Dunlosky et al., 2013). Several previous studies have begun testing this hypothesis, often using the Motivated Strategies for Learning Questionnaire to measure learning strategy uptake (MSLQ; Pintrich et al., 1991). The MSLQ subscales of critical thinking, elaboration, and metacognitive self-regulation are typically classified as deep strategies, whereas the subscale of rehearsal is typically classified as a shallow strategy.

Senko et al. (2011) reviewed the studies that reported associations between goals and strategies. Among the 14 studies they found, the average association between performance-approach goals and surface strategies was $r = 0.20$. Among 13 studies, the association between mastery-approach goals and surface learning was also $r = 0.20$, indicating that perhaps both mastery-approach and performance-approach goals were approximately equally predictive of surface learning. However, Senko et al. were reporting average raw correlations, and not necessarily values derived from a multiple regression analysis, so this analysis cannot determine the extent to which these types of goals predict surface learning over-and-above the other. A multiple regression approach is particularly warranted in the context of goal orientation theory, as mastery-approach and performance-approach goals are often positively correlated with one-another (Harackiewicz et al., 2002; e.g,. Mouratidis et al., 2018). Therefore, in the following paragraphs we summarize a subset of the studies examining the links between goal orientations and learning strategies to parse out findings from their regression models and some of the common methodological and sample-related choices made by these researchers.

Liem et al. (2008) represents one particularly rigorous test of the hypothesized associations between goal orientations and learning strategies in a large (N = 1475) nationally representative of Singaporean ninth graders with regard to their English classes. They found that mastery-approach goal orientations predicted increased reporting of using both deep and surface learning strategies as measured by the MSLQ. Interestingly, performance-approach goal orientations were linked only to deep learning, not shallow learning, while performance-avoidance goals were linked to surface learning and not deep learning. Senko and Miles (2008) recruited 260 undergraduate students from a general psychology course and found that mastery-approach goals, but not performance-approach goals, were predictive of both higher deep and surface learning strategies in a multiple regression SEM model. A similar result was found by Ranellucci et al. (2013) in a small sample (N = 73) of undergraduate students, where the use of deep and shallow processing strategies was coded by experimenters from participants' think-aloud com-

ments. In another sample of undergraduate students (N = 256) using the MSLQ, Senko et al. (2013) found somewhat contradicting results to Senko and Miles (2008), with mastery-approach goals predicting deep, but not surface learning. Conversely, performance-approach goals predicted only surface learning.

The present study grew out of a pre-registered replication of Muis and Franco (2009) for the Center for Open Science's contribution to DARPA's Systematizing Confidence in Open Research and Evidence (SCORE) project (Alipourfard et al., 2021). Similar to the studies discussed above, Muis and Franco's study concerns associations between motivational variables (goal orientations as measured by the Achievement Goal Questionnaire; Elliot and McGregor, 2001) and study strategies representing varying depth of processing (elaboration, critical thinking, and rehearsal; measured by the MSLQ). In particular, Muis and Franco 2009) found that mastery approach goals were predictive of elaboration (a deep strategy), critical thinking (deep), metacognitive self-regulation (deep) and rehearsal (shallow) strategies.

As shown in Figure 1, they found that performance-approach goals predicted increased critical thinking (deep) and decreased metacognitive self-regulation (deep). Performance-avoidance goals were linked with increased rehearsal (shallow). Mastery-avoidance goals were linked with decreased elaboration (deep).

Open Questions Regarding Goal Orientations and Learning Strategies

Altogether, it appears goal orientations and learning strategies are associated, the universality and degree of these associations is unclear. There are several possible reasons for the seemingly contradictory results in this literature area. One area of limitation with regard to pre-existing studies of these associations is that they have been primarily conducted in the context of single institutions or single classrooms (e.g., Muis and Franco, 2009), with an over-representation of students in psychology courses. Although Liem et al.'s (2008) national study of Singapore presents a notable exception to this generalization, we cannot be assured that their results would transfer one-to-one to a North American context.

Heterogeneity and Hidden Moderators

As a result of the limited contextual variability inherent in these studies, it is impossible to disentangle the impacts of potential (hidden) moderators, such as course type, assignment characteristics (e.g., test-based versus essay-based assignments), and individual level variables, such as race and gender. All of these variables have theoretically plausible reasons for moderating the relationships between goal orientations and learning strategies as they are already implicated in motivation theories with regard to their main effects on focal variables. Per Broekkamp and Van Hout-Wolters (2006), students' learning strategies vary depending on academic discipline, type of assignment, and even across individual courses. Even the mere expectation of test format shapes the mnemonic encoding strategies that students engage while studying (Rivers and Dunlosky, 2020).

Cultural differences in learning strategy use have also been found, in terms of both institutional culture and student demographic characteristics. For example, both school-wide institutional and classroom specific culture shape the perception and adoption of mastery goals at the student level (Deemer, 2004). Student gender and ethnicity is also associated with different patterns of motivation variables and self-regulated learning strategies (e.g., Li, 2019; Psaltou-Joycey, 2008). D'Lima et al. (2014) found that female students were more mastery orientated than male students, who conversely tended to be more performance oriented. In terms of learning strategy use, Purdie et al.'s (Purdie et al., 1996) study of Australian students, Japanese students in Japan, and Japanese students in Australia shows that cultural background influences the types of learning strategies students adopt, with Japanese students endorsing higher levels of memorization strategies and lower levels of goal setting strategies. Furthermore, Japanese students studying in Australia were found to be influenced by both their experience in Japanese and Australian school systems, falling between the Australian and Japanese in Japan students in terms of study strategies.

Extant research shows that both motivational and self-regulated learning processes are driven by a variety of variables related to the task at hand, goals being pursued, and student backgrounds. Although main effects on motivational variables and the use of study strategies have been found, less is known regarding the links between goal orientations and study strategies. In particular we were interested in whether these previously reported associations are replicable, and whether they are moderated by any of these educationally relevant variables.

Lack of Replications

Another source of the difficulty in evaluating this research area is the lack of published pre-registered replication attempts. Although open science is not necessarily a new idea in other research domains (e.g., physics; Hecker, 2017), within education research open scientific practices, such as pre-registration of hypotheses, open data, and open analysis scripts are not yet the

norm (van der Zee and Reich, 2018). Indeed, of the studies we surveyed in our literature review, none of these studies were pre-registered, nor did we find any openly accessible datasets. Given the high researcher degrees of freedom inherent in structural equation modelling (Seaman and Weber, 2015) and survey research generally, different patterns of results could stem from a variety of measurement, analytic, and reporting decisions that are essentially unverifiable in the absence of open scientific practices (see also Crede and Harms, 2019). Although not a panacea in and of themselves, pre-registered replication studies attempt to ameliorate some of these issues by laying out the hypotheses, analysis plan, and participant sampling decisions ahead of time (Plucker and Makel, 2021).

The Present Study

Although, given the study design used by Muis and Franco (and, thus, the present study), no causal conclusions can be made, these associations have key mechanistic implications for the way in which motivational variables may operate—particularly in terms of how achievement motivation may work to increase classroom performance. While maintaining as much fidelity to the original study as possible, the present research effort builds on Muis and Franco's work in several key ways. Whereas Muis and Franco's study was limited to 201 students recruited from a single educational psychology course, the present study recruited 978 undergraduate students across numerous courses and universities in the United States and Canada.

While most of the previous work in this research area has drawn student participants from single universities (and often single classes within these universities), the relative breadth of our sample allows for the comparison of these associations across general types of classes (i.e., STEM, Social Sciences, and Humanities) outside of the context of a single university. Should associations hold up under such relatively wide heterogeneity of university experience, we can be relatively confident that the associations should generalize to other contexts within the United States and Canada.

Given the differences in the type of learning materials taught across these broad divisions of the academy, we were interested in whether the associations between motivational beliefs and learning strategies were invariant across many different types of classes. Furthermore, our relatively large sample allows us to test measurement invariance across other theoretically interesting sub-groups, including underrepresented minority populations and gender, where issues concerning achievement disparities (Museus and Liverman, 2010) and structural racism/sexism (McGee, 2020) have been identified previously in both the American (Hurtado and Ruiz Alvarado, 2012) and Canadian contexts (Henry et al., 2017). Together, the qualities of this sample and the pre-registered analysis and data collection plan allows us to build on previous work concerning the associations between study strategies and goal orientations in an incremental, yet rigorous manner.

Therefore, in the present study, our two main research questions are: (1) How is each type of achievement goal uniquely related to study strategies? (2) Do the relationships vary across course types and student sub-groups (underrepresented minority status, gender)? For the first RQ, we drew upon the results of Muis and Franco (2009; see Figure 1) and predicted that each of the four study strategies of interest (Elaboration, Critical Thinking, Metacognitive Self-Regulation, and Rehearsal) would be predicted by a unique set of goal orientations. For the second RQ, we did not have specific pre-registered predictions. We present these analyses as exploratory and include them because theoretically one might expect the associations uncovered by our model to differ by student background.

Methods

The materials for this study, including the pre-registration, power analysis, and deidentified data set are available on the Open Science Framework's data repository (https://osf.io/e7tw2/). The pre-registration is located here: https://osf.io/x9kda; the claims tested in this paper are located in the exploratory section of this document, because they were registered in addition to the focal claim selected by the SCORE program.

Participants

Participants were recruited from Prolific (prolific.co), an online survey recruitment platform during October 2020. Participants were required to be currently enrolled in an undergraduate, community, or technical college and a resident of the United States or Canada as determined by the Prolific eligibility screener. We note that the pre-registration listed only undergraduate and community college students, but technical college students were allowed to register for the study due to concerns about sample-size limitations. This decision was made prior to any data being collected.

The United States and Canada were chosen for this survey, because they share the same language and general school calendar. Altogether, 1,018 students were recruited to the survey. Forty students failed at least one attention check. Per our pre-registered exclusion criteria, these students were removed from the survey for the purpose of our analyses, for a final sample size of N = 978. This sample size was chosen in accordance

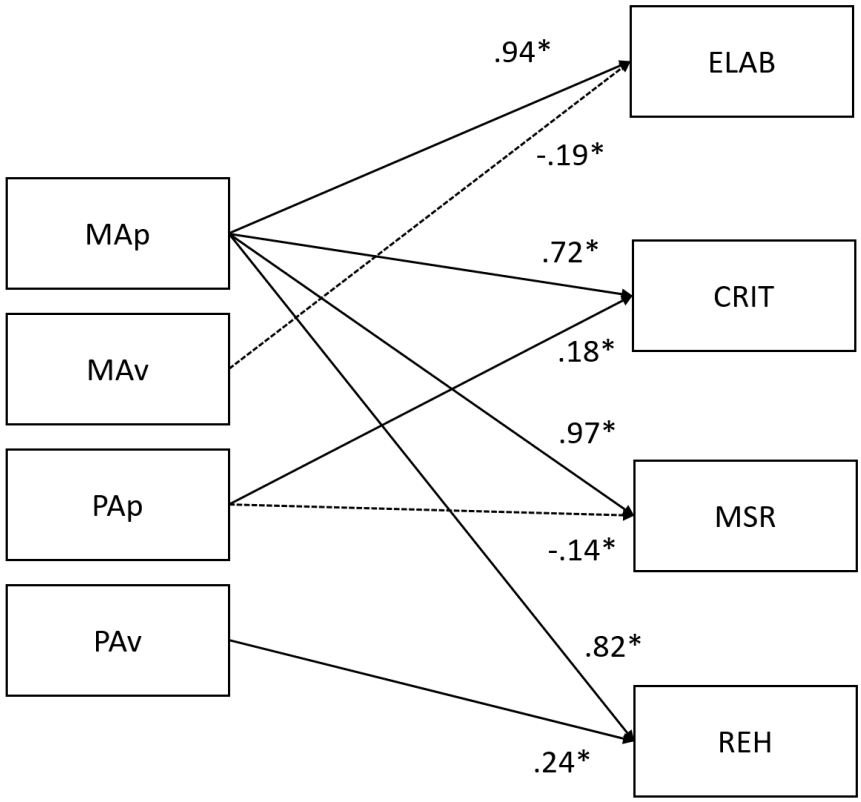

Figure 1

Structural Equation Model Diagram Showing Results of Muis and Franco (2009)

*Note. Solid lines indicate significant positive associations. Dashed lines represent significant negative associations. The following latent variables are depicted: Mastery Approach (MAp), Mastery Avoidance (MAv), Performance Approach (PAp), Performance Avoidance (PAv), Elaboration (ELAB), Critical Thinking (CRIT), Metacognitive Self-Regulation (MSR) and Rehearsal (REH). Significance represented as follows: * $p < .05$ (no further cutoffs were reported in original paper).*

with a power analysis conducted by the Center for Open Science in order to successfully detect a standardized structural equation modeling coefficient of 0.41, with 90 percent power.

The demographics of this final sample are as follows. The mean age was 22.22 years (Median = 21.00, 2 Missing). 576 (59%) participants identified as female, 398 (40%) identified as male, and 4 (<1%) answered "Prefer Not to Say" or had missing data. 863 (88%) participants were current residents of the United States, while 108 (11%) were residents of Canada (7 or <1% Missing). Note that the gender and nationality demographic questions were pre-administered by the Prolific screener, and thus the research team was not able to ask finer-grained questions concerning these identities. In terms of race or ethnicity, participants were asked "Which of the following best describes you?" and 274 (28%) answered "Asian or Pacific Islander;" 69 (7%) "Black or African American/Canadian;" 88 (9%) "Hispanic or Latino;" 4 (<1%) "Native American, First Nations, or Alaskan Native;" 474 (48%) "White or Caucasian;" 61 (6%) "Multiracial or Biracial;" and 8 (<1%) "a race/ethnicity not described here." Thus, 230 students were part of historically underrepresented groups in higher education.

Each participant chose a single course at the beginning of the survey to answer the questions in response to, they were asked to categorize this course in terms of the general area of study. 509 (52%) students answered the survey regarding a "STEM: Natural Sciences, Mathematics, Computer Science, Engineering" course, while 243 (25%) answered with regard to a "Social Sciences (e.g., Economics, Business, Psychology, Anthropology)" course, and 226 (23%) chose a "Humanities or Lan-

guage Arts (e.g., English, Comparative Literature, Philosophy)" course. In terms of assignments mentioned, 673 participants mentioned completing quizzes, exams, or tests. The remaining 305 participants mentioned a mix of group projects, essays, or alternative assessment types.

Participants were paid $1.50 each, or approximately $9.55 per hour, given that the survey took participants a little under nine-and-a-half minutes to complete on average. This research was conducted in accordance with the oversight of the University of Texas at Austin's Institutional Review Board and also the Naval Information Warfare Center Pacific, Human Research Protection Office.

Materials

Four scales were administered for the purpose of this survey. One scale was ad-hoc in construction (Memorability), while three others were drawn without modification from the research literature.

Goal Orientations

The 12-item Achievement Goal Questionnaire published by Elliot and McGregor (2001) was used to measure goal orientations. Four subscales, mastery approach, mastery avoidance, performance approach, and performance avoidance, were each measured using three items. An example performance approach item is "It is important for me to do better than other students," indicative of being performance-oriented (motivated by grades) and also approach-oriented (seeking to outperform). An example mastery avoidance item is "Sometimes I'm afraid that I may not understand the content of this class as thoroughly as I'd like." It is mastery-oriented because it involves wanting to learn the content (as opposed to being grade/performance oriented), and it is avoidance-oriented because it reflects a fear of doing poorly. Items were measured using Likert questions ranging from "Not at all true of me" to "Very true of me."

Self-Regulated Learning and Metacognitive Strategies

A subset of the items from the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991). These subscales were chosen to measure the various learning strategies students used and to map onto the previous work of Muis and Franco (2009). The items used in this survey were drawn from four subscales, Elaboration (6 items), Critical Thinking (5 items), Rehearsal (4 items) and Metacognitive Self-Regulation (12 items) drawn from the larger MSLQ inventory.

The rehearsal scale measures the use of rote (shallow) learning strategies (e.g., "When I study for this class, I practice saying the material to myself over and over"). Elaboration measures deeper processing study strategies (e.g., "When I study for this course, I write brief summaries of the main ideas from the readings and the concepts from the lectures"). The critical thinking subscale is indicative of a student's self-reported use of analytic thinking and drawing connections between disparate sets of knowledge (e.g., "Whenever I read or hear an assertion or conclusion in this class, I think about possible alternatives."). Metacognitive self-regulation measures the extent to which students engage in self-monitoring, planning, and metacognitive control (e.g., "When I become confused about something I'm reading for this class, I go back and try to figure it out."). These questions were answered on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree."

Memorability

This scale was composed of three questions aimed at determining the extent to which students felt they could accurately recall their course and studying activities. They were asked two Likert questions: "How well can you remember what you did in order to prepare for these recent assignments/quizzes/exams?" and also "If a friend asked you to recount to them how you studied for these assignments/quizzes/exams, how confident are you in your ability to give them an accurate description of what you did?", both on a scale of one-to-seven. Additionally, they answered the question: "How long ago were the assignments you were thinking about?" To this question, participants could answer either: "Last week," "Two weeks ago," "Within the last month (3-4 weeks ago)," and "More than a month ago." Given that memory is known to degrade over time, this last question was asked to indirectly measure memory strength or recallability of the course assignments and student study strategies. This scale was not pivotal to the hypotheses of our study nor was it used by Muis and Franco (2009), rather these questions were included in order to make sure that there were no systematic differences in memorability across different groups of respondents (see pre-registration for further discussion for why this scale was included in the study).

Disciplined-Focused Epistemic Beliefs

Although these were not a focal part of the replication, we also administered the 18-question, four-factor Disciplined-Focused Epistemological Beliefs Questionnaire as developed by Hofer (2000). The four previously established factors are Certainty, Justification

(Personal), Source of Knowledge (Authority), and Attainment of Truth. Muis and Franco (2009) used a longer 27-item scale that adds an additional nine items to measure the same four constructs; we used only the validated 18 items.

Procedure

Students were recruited to the survey via the Prolific.co online recruitment platform. They read a consent form and if they consented to proceed were advanced further into the Qualtrics survey platform. Students were then asked which course they would like to answer the survey with regard to:

> Please think of one single ongoing course that you are currently enrolled in. You will need to answer all of the following survey questions with regard to this course. Please write the name of the course below (e.g., Intro Psych, Calculus 1, Geographic Information Systems)

They were then asked to think about the course and describe its assignments in as much detail as possible in an open-ended text box:

> You have chosen an ongoing course that you are currently enrolled in, [Course Name].
>
> Please think of the major assignments, such as essays, quizzes, or exams, that you have completed in this class. Describe the assignments in the text box below, and any other details that come to your mind while writing. The more detail, the better.
>
> You may spend as long as you choose answering this question, but the next page button will appear only after two minutes of writing.

Students were allowed to proceed after two minutes of contemplation and writing. Although clearly not a one-to-one substitute, this activity was inspired by Muis and Franco's (2009) reflection activity, wherein students reviewed their previous assignments in an educational psychology course before answering questions regarding their goal orientations, epistemic beliefs, and self-regulated learning strategies. They were then asked three questions regarding how memorable the course was as described in the materials section above. After answering these memorability questions, they were presented with the goal orientation (AGQ), epistemic beliefs (DFEBQ), and self-regulated learning (MSLQ) scales in a random order. For example, some students answered AQG, then DFEBQ, and MSLQ, while others would have answered the MSLQ first, then AGQ, and DFEBQ. Within each scale, the items were presented in a random order per participant.

Two attention checks were placed randomly within these survey questions. One attention check asked participants to answer "Strongly Agree," while the other asked students to respond "Strongly Disagree." After completing the four scales of interest, each participant categorized the chosen course as STEM, Social Sciences, or Humanities in nature. Then they reported whether they had received grades for the course assignments already (either "Yes, I have received grades for at least some of these assignment(s)" or "No, no grades have been received"). And reported their race or ethnicity as described above. Upon completing the survey, students were thanked for their time and returned to the Prolific platform.

Results

The primary goal of the replication attempt was to analyze the associations between the four scales drawn from the Achievement Goal Questionnaire and the four subscales drawn from the Motivated Strategies for Learning Questionnaire.

Confirmatory: Associations between Goal Orientations and Study Strategies

In this pre-registered analysis, a structural equation model was built using seven latent variables, one for three of the subscales of the MSLQ (Rehearsal, Elaboration, Metacognitive Self-Regulation, Critical Thinking) and four subscales of the AGQ (Mastery Approach, Mastery Avoidance, Performance Approach, Performance Avoidance). The Likert data was entered in as ordered items, and the model was fit using the weighted least squares with robust standard errors (WLSMV) using the R version 4.0.0 (R Core Team, 2020) package Lavaan version 0.6-9 (Rosseel, 2012). The model fit moderately well, RMSEA = 0.073, CFI = 0.962, TLI = 0.958, though given the sample size it was not surprising that the scaled chi-squared test was significant $\chi^2(674) = 4189.275$, $p < .001$ (see Shi et al., 2019; Xia and Yang, 2019). As shown in Figure 2, we found that the four study strategies latent variables were predicted by endorsement of different goal orientations.

Assessment of Stability of Associations Across Varying Contexts

In Table 1, we present the average scale score of each dependent variable, by moderator sub-group. This table shows that mean strategy endorsement is very similar across the different groups and there are only a few

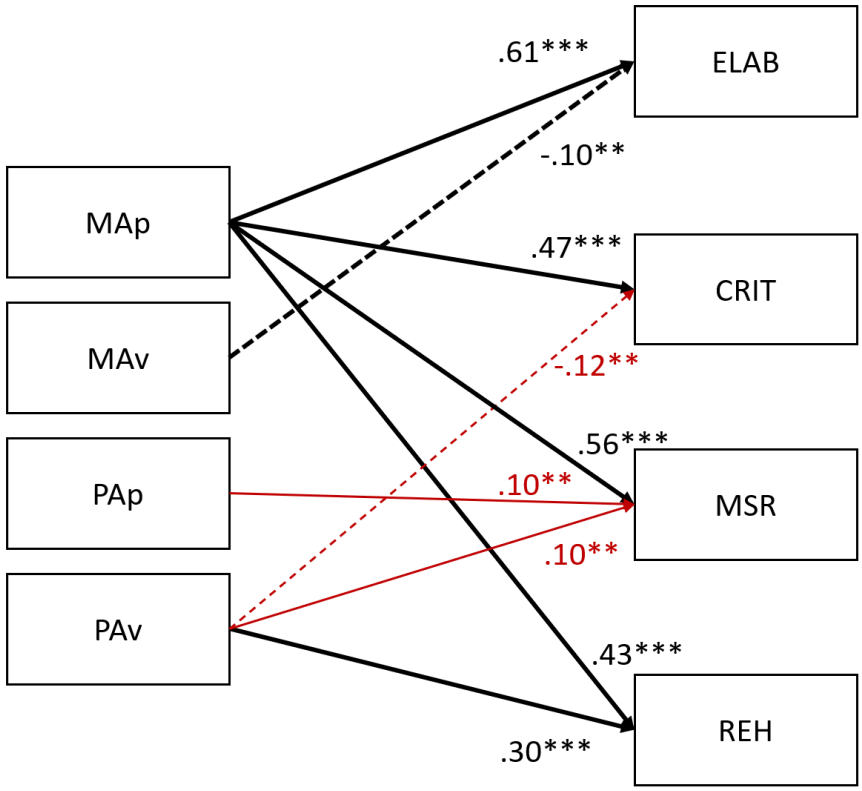

Figure 2

Structural Equation Model Diagram Showing Replication Results.

*Note. Thick black lines indicate associations that replicated successfully. Thin red lines indicate significant associations in the present sample that were not significant in Muis and Franco (2009) or did not replicate their findings in terms of sign. Solid lines indicate significant positive associations. Dashed lines represent significant. * $p < .05$, ** $p < .01$, *** $p < .001$. Standardized coefficients shown.*

moderate-to-large differences in mean strategy endorsement (e.g., STEM students reported less critical thinking than Humanities students and students in exam-based classes reported less critical thinking than those describing non-exam-based classes). In light of the context-dependent nature of motivation and learning strategies, we sought to determine if the strength of the associations between goal orientations and learning strategies were moderated by the characteristics of the students and the courses they were taking. That is, regardless of mean endorsement, do the associations between goal orientations and learning strategies vary across moderator subgroups.

Throughout our moderation tests, we used a multiple groups analysis approach. In this approach, two models are estimated. In the simpler model, regression coefficients are held constant across groups. In the second, more complex model, regression coefficients are allowed to vary by group. Then, a scaled chi-squared difference test is used to compare the two models. If there is a significantly better fit of the more complex model (allowing regression coefficients to differ by group), this indicates a significant omnibus test of the moderation of the associations (Cortina et al., 2021).

In a departure from earlier models reported in this paper, we did not enter the metacognitive self-regulation factor as a dependent variable and thus did not estimate any regression coefficients between the four goal orientation variables and the MSR factor. This analytic decision was made due to issues with model identifiability (certain subgroup models could not be estimated). It was found that the MSR factor caused issues with estimation ostensibly due to its poor factor structure (see exploratory factor analysis output in the online materials). We pre-registered the analysis of differences by course type. The other moderation analyses

Table 1

Mean and standard deviation strategy endorsement by subgroup.

Moderator	Subgroup	N	Elaboration	Critical Thinking	Metacognitive Self-Regulation	Rehearsal
Course Type	STEM	509	4.86 (1.10)	3.89 (1.34)	4.61 (0.84)	4.37 (1.36)
	Social Sciences	243	4.96 (1.17)	4.17 (1.25)	4.50 (0.95)	4.26 (1.47)
	Humanities	226	5.01 (1.09)	4.57 (1.25)	4.50 (0.96)	4.16 (1.68)
Assignment Type	Exams	706	4.89 (1.10)	3.97 (1.34)	4.55 (0.87)	4.36 (1.44)
	Non-Exams	272	4.99 (1.16)	4.50 (1.21)	4.57 (0.96)	4.13 (1.53)
Race and Ethnicity	AAPI	274	4.70 (1.02)	3.91 (1.24)	4.49 (0.83)	4.23 (1.37)
	URM	230	4.92 (1.23)	4.27 (1.33)	4.56 (0.95)	4.39 (1.55)
	White	474	5.05 (1.10)	4.16 (1.36)	4.59 (0.90)	4.28 (1.49)
Gender	Men	398	4.82 (1.09)	4.22 (1.27)	4.55 (0.87)	4.30 (1.37)
	Women	576	4.98 (1.14)	4.05 (1.36)	4.56 (0.91)	4.29 (1.53)

Note. STEM = Science, technology, engineering, and mathematics; AAPI = Asian American and Pacific Islander; URM = underrepresented minority (Black, Indigenous, Hispanic/Latino students). Subgroup Ns may not add up to the total sample size of N = 978 due to missingness.

were exploratory.

Confirmatory: Moderation by Course Type

Differences in regression coefficients (i.e., the associations between goal orientations and study strategies) were tested across the three different types of course classifications, STEM, Social Science, and Humanities. We did not find evidence in support of differing regression coefficients across the course types, $\chi^2(24) = 32.75$, p = .109. This result indicates that the associations of interest between the latent variables were relatively stable across the different course types.

Exploratory: Moderation by Assignment Type

Participant's written text description of the assignments they completed as part of their focal course were analyzed for whether they included mentions to quizzes, exams, or tests. Of the 978 participants, 706 reported having quizzes, tests, or exams as part of their course assignments, while 272 made no mention of these assignment types. This variable was coded as a binary variable (either included mention to quizzes/exams/tests or did not). The same approach as described above was used to test whether the regression coefficients differed between the two groups. Here, significant evidence of moderation was found, $\chi^2(12) = 27.45$, p = .007.

Post-hoc comparison revealed that constraining the following coefficients significantly worsened model fit, thus indicating that they should be estimated separately. The coefficient of Mastery Approach predicting Critical Thinking was significantly higher in the non-exam group (β = 0.60) than the exam group (β = .44), $\chi^2(1) = 4.83$, p = .03. Similar results were found with regard to Mastery Avoidance predicting Elaboration with a stronger association in the non-exam (β = 0.70) than exam group (β = 0.56), $\chi^2(1) = 5.04$, p = .02. There was also a stronger relationship between Performance Avoidance and Elaboration in the non-exam group (β = 0.20) than the exam group (β = −.01), $\chi^2(1) = 5.85$, p = .02.

Exploratory: Moderation by Race and Ethnicity

Historically in the United States and Canadian higher education systems, Hispanic/Latinx, Black and Indigenous students have been underrepresented (Henry et al., 2017; Hurtado and Ruiz Alvarado, 2012). Numerous studies have also shown differential educational beliefs among these student populations (e.g., Fong et al., 2019). Thus, we were interested in determining if the associations between goal orientations and study strategies were moderated by underrepresented minority status (Asian and Pacific Islanders, URM, White). Here, underrepresented minority denotes students coming from Black, Indigenous, Hispanic/Latino backgrounds. Although it would be preferable to split our analysis into smaller groups (i.e., Asian, Black, Hispanic/Latinx, Native American and Indigenous, and White), this was not

tenable with the present sample size.

Similar to the moderation by course type analysis presented above, we created one model where the regression coefficients were held constant across the three groups (AAPI, URM, White) and one model where the regression coefficients were freely estimated across the three groups. We then compared the relative increase in model fit with the freely estimated model in comparison to the held-constant model, and did not find statistically significant evidence of moderation by URM status, $\chi^2(24) = 23.90$, $p = .467$.

Exploratory: Moderation by Gender

A similar approach was used to test whether the regression coefficients were moderated by gender as the course type and URM analyses. We did not find evidence that model fit was significantly better when allowing regression coefficients to differ by gender (Male or Female), $\chi^2(12) = 9.83$, $p = .631$. This analysis was conducted using a sample size of $N = 974$, as four participants did not report their gender.

Exploratory: Replication Attempt of Expanded Model

In addition to our specific interest in the relationship between motivational beliefs and study strategies, we also attempted to replicate Muis and Franco's (2009) larger structural equation model (a subset of which is shown in Figure 1 of this paper; see Figure 1 in Muis & Franco for the entire model), however, despite recruiting 777 more participants in our sample, we did not find that the model outlined in their paper was identifiable. Even after making changes to improve model fit (such as removing the metacognitive self-regulation factor), we did not find success in finding a model with adequate fit metrics in terms of CFI and RMSEA.

Issues with model identifiability seem to stem from the fact that the latent variables derived from the DFEBQ (Certainty, Personal Justification, Source Authority, and Attainability of Truth) were highly correlated with one-another (absolute values of r ranging between .63 and .91; see Table 2). A confirmatory factor analysis of the three measures (DFEBQ, MSLQ, and AGQ; 12 total subscales) showed moderately acceptable fit, CFI = 0.944, TLI = 0.939, RMSEA = 0.077, $\chi^2(1473) = 10111.05$, $p < .001$.

Although the CFA fit metrics were passable, issues with the DFEBQ scale were further confirmed by an exploratory factor analysis of items from the three scales (AGQ, DFEBQ, and MSLQ). Although a priori theory suggests there should be twelve sub-factors across these three measures, parallel analysis (Horn, 1965) using the command "fa.parallel" in the psych R package (Revelle, 2018) only recommended ten factors, and when the analysis of these ten factors was explored, this analysis found that a majority of items (16 out of 18) from the DFEBQ loaded on a single factor. Even when an exploratory factor analysis with 12 factors was computed, it did not lead to items loading where theory would predict, as items from the DFEBQ scale still tended to cluster together with one another rather than splitting into four separate factors.

It would be plausible to suggest that the slightly abbreviated version of the DFEBQ scale used in this study may have contributed to issues with model fit, however, we do not believe this to be the case. As shown in Table 2, the Cronbach's alphas for the Justification (Personal) and Source (Authority) were lower than .70, indicating less than ideal reliability (Tavakol and Dennick, 2011). Though an abbreviated version of the DFEBQ was used in comparison to that used by Muis and Franco (2009), the Cronbach's alphas in the present sample were not substantially worse than reported by Muis and Franco. In fact, in three out of four scales the reliability coefficients in our sample were higher, with the sole exception being Justification (Personal). Although generally Cronbach's alpha increases with scale length, the counterintuitive finding here might be explained by the fact that the shorter version used in this present study only consisted of the 18 items previously validated by Hofer (2000), while the nine additional items were not recommended for use by Hofer.

Additionally, our sample was recruited across many different classes and (ostensibly) universities. Muis and Franco recruited participants from a single Educational Psychology class, which may have aided in identifying the factor structure given the self-ascribed "discipline-focused" nature of the DFEBQ. This hypothesis is somewhat supported by the distributions of sum scores across the three disciplines (STEM, Social Sciences, Humanities), which shows quite distinct patterns in epistemic beliefs, with several pairwise comparisons indicating differences in sum scores equivalent to Cohen's ds of 0.8 or greater between STEM and Humanities/Social Sciences. Thus, it seems that epistemic beliefs are at least partially dependent on the specific discipline a student is engaging with.

Discussion

This paper adds to the growing body of research on the associations between self-reported goal orientations and learning strategies, adding further evidence to the theory that the relationship between increased self-report motivation and achievement may be partially mediated by the use of better learning strategies. Altogether, eleven of the sixteen (69%) associations tested

Table 2

Correlations between latent variables, reliabilities, and sum scores.

	JUST	ATTAIN	CERT	AUTH	MAv	MAp	PAp	PAv	REH	ELAB	CRIT	Alpha	Mean (SD)
JUST	1.00											.63	3.9 (1.3)
ATTAIN	−0.52*	1.00										.74	4.9 (1.4)
CERT	0.65*	−0.77*	1.00									.88	3.8 (1.4)
AUTH	−0.55*	0.83*	−0.83*	1.00								.69	4.7 (1.2)
MAv	−0.07	0.20*	−0.16*	0.26*	1.00							.81	4.7 (1.6)
MAp	−0.03	0.21*	−0.02	0.10	0.39*	1.00						.86	5.4 (1.4)
PAp	−0.03	0.15*	−0.18*	0.21*	0.11*	0.32*	1.00					.90	4.5 (1.7)
PAv	−0.11*	0.24*	−0.18*	0.35*	0.32*	0.01	0.31*	1.00				.78	5.5 (1.4)
REH	−0.05	0.34*	−0.18*	0.27*	0.25*	0.43*	0.23*	0.30*	1.00			.75	4.3 (1.5)
ELAB	0.14*	0.11	0.17*	0.01	0.16*	0.57*	0.20*	0.03	0.52*	1.00		.75	4.9 (1.1)
CRIT	0.51*	−0.12*	0.34*	−0.29*	0.09	0.47*	0.15*	−0.12*	0.28*	0.70*	1.00	.80	4.1 (1.3)
MSR	0.05	−0.32*	0.12*	−0.25*	−0.29*	−0.60*	−0.31*	−0.14*	−0.71*	−0.79*	−0.62*	.77	4.5 (0.9)

Note. Latent variable correlations derived from a confirmatory factor analysis performed in Lavaan using item groupings derived from previously published scales. Asterisk indicates significant correlation at $p < .001$. Mean sum scores calculated after reverse scoring appropriate items.

in both the present paper and Muis and Franco (2009) replicated in terms of statistical significance and direction of association.[1] Though not a perfect replication rate, all associations larger than .20 in Muis and Franco's original study successfully replicated. If we limit our analysis to only significant associations shown by Muis and Franco, 75% replicated (six out of eight).

Like Muis and Franco (2009), we found that the links between mastery-approach orientation and three strategies of interest (elaboration, rehearsal, metacognitive self-regulation and critical thinking) were all strong, statistically significant, and positive in nature. The link between mastery-approach and elaboration was the strongest, followed by metacognitive self-regulation, critical-thinking, and lastly rehearsal, suggesting that mastery-approach slightly favors deeper study strategies. This suggests that mastery-approach goals are related to increased overall effort. In other words, students who report mastery-approach goals engage in more shallow strategies and more deep strategies than those not endorsing such a goal orientation.

Although we tested for heterogeneity within our sample, we found only one significant moderator of these associations: assignment type. Taken together, it appears that the associations between goal orientations and learning strategies may be stronger when courses use more open-ended evaluation (e.g., essays, lab reports, projects, presentations) than exam-based evaluation. Given the replication of earlier results and relative lack of heterogeneity within our sample, this suggests that there are relatively consistent associations between these measures, even when assessed in different populations and different time periods. However, we do note that on average the coefficients in our model were approximately half as large as their counterparts in Muis and Franco, which suggests the original estimates may have been inflated by the design or more homogeneous context of the original study.

Limitations

One of the critical differences in methodology between the present paper and that of Muis and Franco (2009) is that our survey was conducted during the midst of the COVID-19 pandemic (October 2020) in countries where the virus was prevalent. This project timeline was a stipulation of the funding agency and outside of the research team's control. As a result, many of the colleges and universities in the sampled countries were partially or even entirely online (Anderson, 2020). Online or hybrid coursework presents interesting motivational situations that may not be entirely applicable to in-person schoolwork (Chen and Jang, 2010). Although this aspect of the study seems like a limitation on its face, it may actually be a stronger test of the replicability of the association. Considering that most of the associations (75%) found by Muis and Franco (2009) were

replicated in the present study, it seems that these associations may be robust to changes in educational (i.e., across different courses and institutions) and social contexts (i.e., pre-pandemic versus during pandemic).

There were however small differences between our findings and that of prior studies: We did not find any significant associations between performance-approach goals and critical thinking. In contrast, others have shown small significant correlation between performance-approach and critical thinking (Muis and Franco, 2009; Vrugt and Oort, 2008). That being said, given the array of disparate findings concerning the associations between goal orientations and learning strategies reviewed earlier in this paper, it is likely that these associations are somewhat scale/measure specific. For this reason, more research needs to be done to understand where existing scales do and do not overlap and whether they are truly measuring the same constructs.

One limitation of the present study is that we could not collect achievement measures (i.e., grades) from the students in this survey. Because these students were recruited from an online participant pool from a variety of educational institutions during the middle of the semester, final grades would not have been available to the students. Thus, due to the single-session nature of our survey, we would have had to use anticipated grades. Self-reported grades (i.e., students reporting the actual grades or grade point averages they receive) are known to have relatively poor validity (Kuncel et al., 2005), and self-reported anticipated grades would likely be subject to similar, if not worse, issues. Additionally, not all Canadian institutions use the same grade-point scale as universities in the United States, which would have further complicated issues of measurement. For this reason, we chose not to collect achievement-related measures. Past studies, such as Muis and Franco (2009) have found correlations between learning strategies and true final grades, but future work should work to determine the extent to which these correlations are generalizable across institutions and individuals. Moreover, it would be important to understand what final grades truly represent. For example, the extent to which grades are based on participation and group work may affect the relationship between individual learning strategies and grades. It is also important to consider efficiency: two students may attain an A-grade in very different ways—one could achieve a high grade by studying 80 hours a week using relatively shallow strategies or one could achieve a high grade by studying far fewer hours but using relatively deeper strategies.

Future Directions

As shown in Table 2 and as discussed earlier in this paper, many of the subscales we drew from the published literature were not as reliable and distinct from one-another as one might expect (this is a known issue with measuring epistemic beliefs in particular; see DeBacker et al., 2008; see also generally Flake and Fried, 2020). While the goal orientation measures showed relatively high reliability and the expected structure, the items/subscales from the disciplined-focused epistemic belief questionnaire and motivated strategies for learning questionnaire overlapped with one-another, raising questions of discriminant validity and ultimately causing issues with model fit.

Future researchers might consider moving away from the Motivated Strategies for Learning Questionnaire as the primary measure of strategy use, as this questionnaire does not measure the use of several effective study strategies now widely regarded as highly effective. Specifically the MSLQ, does not ask students about retrieval practice or spacing, the use of which is known to be associated with student achievement (Carvalho et al., 2020; Dunlosky et al., 2013). Additionally, the present study showed relatively poor discriminant validity (particularly with regard to the metacognitive self-regulation factor). Thus, it seems that an updated measure of learning strategy use would be a productive next step. An improved scale could help facilitate the development of more valid, reliable studies and help to integrate research on student beliefs concerning learning strategies with research on the broader cognitive science of learning and memory (see also the growing "desirable difficulties" research in the vein of Bjork and Bjork, 1992). Research should also continue to build on the present correlational study through the continued use of experimental intervention studies to validate the hypothesized causal pathways stemming from this study and others like it.

In order to better the measurement of key educational constructs, we have attempted to follow many of the best practices prescribed by recent work in open science. Transparency was maintained through the preregistration of analytic choices and the demarcation of which analyses were exploratory in nature. The data from this study has also been made available on the Open Science Framework in order to aid future researchers in the study of motivation science and learning strategies. We also encourage researchers interested in conducting replication of studies using SEM methods to be especially careful in terms of choosing studies that use properly validated scales and to be aware of the possibility that their models may not properly converge in a new sample (though potentially advances in Bayesian

SEM will make issues of convergence moot in the near future, once ordinal models are fully supported by these software packages; see Merkle and Rosseel, 2015).

Concluding Comments

In general, our data revealed a relative stability of the associations between achievement goal orientations and study strategies. Not only did we replicate three fourths of the associations reported by Muis and Franco, but the larger sample size and diversity of students in our study also allowed us to test whether the associations were invariant across various sub-groups. We did not find these associations differed significantly across course discipline (STEM, Social Sciences, or Humanities), underrepresented minority status, or gender. We did find evidence that these associations were moderated by assignment type, with essay based classes showing stronger associations between goal orientations and learning strategies. Given that the associations found in this study were largely congruent with the findings of Muis and Franco, and are not significantly different across the subgroups analyzed in the present sample (except for assignment type), the evidence suggests these associations are relatively stable across heterogeneous contexts.

End Note

1. This figure may be an underestimate of replicability as Muis and Franco did not report non-significant coefficients, and given the larger sample size we were able to detect smaller associations than the original study. Thus, theoretically there may be cases where we estimated exactly the same magnitude of association, but it was not significant in the original study and significant in the present study resulting in the somewhat counterintuitive finding of non-replication. There were only two significant associations from the original study that were not significant and in the same direction in the present sample (these were PAp predicting MSR and PAp predicting CRIT).

Author Contact

Correspondence should be addressed to Brendan A. Schuetze, ORCID ID: 0000-0002-5210-6785. B.A.S is now at the University of Potsdam, Campus Golm, Professur Digitale Bildung, Karl-Liebknecht-Str. 24-25, 14476 Potsdam. Email: brendan.schuetze@gmail.com.

Veronica X. Yan, ORCID ID: 000-0002-3988-3184.

Acknowledgements

The authors thank pre-registration editor John Lloyd, and three pre-registration peer reviewers, Jeremy K. Miller, Tina Sundelin, and Zorana Zupan for their thoughtful comments on our study design. We also acknowledge the CoS staff for conducting the power analysis and their help with administering this project.

Conflict of Interest and Funding

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article. This study was supported by SCORE Replication funding to the research team from the Center for Open Science through a grant from the United States Department of Defense, DARPA. The opinions and views expressed in this paper are those of the authors alone and do not necessarily represent the official policy or position of any funding body.

Author Contributions

Both authors applied for funding from the CoS for this project. B.A.S wrote the pre-registration, analyzed the data, and wrote the first draft of the manuscript. V.X.Y. edited the pre-registration, edited the manuscript, and provided guidance throughout the project.

Open Science Practices

This article earned the Preregistration+, Open Data and the Open Materials badge for preregistering the hypothesis and analysis before data collection, and for making the data and materials openly available. It has been verified that the analysis reproduced the results presented in the article. The entire editorial process, including the open reviews, is published in the online supplement.

References

Alipourfard, N., Arendt, B., Benjamin, D. M., Benkler, N., Bishop, M. M., Burstein, M., Bush, M., Caverlee, J., Chen, Y., Clark, C., Dreber, A., Errington, T. M., Fidler, F., Fox, N. W., Frank, A., Fraser, H., Friedman, S., Gelman, B., Gentile, J., & Wu, J. (2021). Systematizing confidence in open research and evidence (SCORE) [preprint]. *SocArXiv*. https://doi.org/10.31235/osf.io/46mnb

Anderson, N. (2020). Examining ways COVID-19 adaptation derailed our assumptions about online learning. *2020 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)*, 67–70. https://doi.org/10.1109/IC3e50159.2020.9288463

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In *From learning processes to cognitive processes: Essays in honor of William K. Estes* (pp. 35–67).

Broekkamp, H., & Van Hout-Wolters, B. H. A. M. (2006). Students' adaptation of study strategies when preparing for classroom tests. *Educational Psychology Review*, *19*(4), 401. https://doi.org/10.1007/s10648-006-9025-0

Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, *64*(8), 1–9.

Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, *18*(1), 26–31. https://doi.org/10.2307/1176007

Carvalho, P. F., Sana, F., & Yan, V. X. (2020). Self-regulated spacing in a massive open online course is related to better learning. *NPJ Science of Learning*, *5*(1), 1–7. https://doi.org/10.1038/s41539-020-0061-1

Chen, K.-C., & Jang, S.-J. (2010). Motivation in online learning: Testing a model of self-determination theory. *Computers in Human Behavior*, *26*(4), 741–752. https://doi.org/10.1016/j.chb.2010.01.011

Cortina, J. M., Markell-Goldstein, H. M., Green, J. P., & Chang, Y. (2021). How are we testing interactions in latent variable models? surging forward or fighting shy? *Organizational Research Methods*, *24*(1), 26–54. https://doi.org/10.1177/1094428119872531

Crede, M., & Harms, P. (2019). Questionable research practices when using confirmatory factor analysis. *Journal of Managerial Psychology*, *34*(1), 18–30. https://doi.org/10.1108/JMP-06-2018-0272

DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The challenge of measuring epistemic beliefs: An analysis of three self-report instruments. *The Journal of Experimental Education*, *76*(3), 281–312. https://doi.org/10.3200/JEXE.76.3.281-314

Deemer, S. (2004). Classroom goal orientation in high school classrooms: Revealing links between teacher beliefs and classroom environments. *Educational Research*, *46*(1), 73–90. https://doi.org/10.1080/0013188042000178836

D'Lima, G. M., Winsler, A., & Kitsantas, A. (2014). Ethnic and gender differences in first-year college students' goal orientation, self-efficacy, and extrinsic and intrinsic motivation. *The Journal of Educational Research*, *107*(5), 341–356. https://doi.org/10.1080/00220671.2013.823366

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Engelschalk, T., Steuer, G., & Dresel, M. (2017). Quantity and quality of motivational regulation among university students. *Educational Psychology*, *37*(9), 1154–1170. https://doi.org/10.1080/01443410.2017.1322177

Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, *3*(4), 456–465. https://doi.org/10.1177/2515245920952393

Fong, C. J., Alejandro, A. J., Krou, M. R., Segovia, J., & Johnston-Ashton, K. (2019). Ya'at'eeh: Race-reimaged belongingness factors, academic outcomes, and goal pursuits among indigenous community college students. *Contemporary Educational Psychology*, *59*, 101805. https://doi.org/10.1016/j.cedpsych.2019.101805

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., Katzaroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal*, *34*(3), 513–543. https://doi.org/10.3102/00028312034003513

Harackiewicz, J. M., Barron, K. E., Carter, S. M., Lehto, A. T., & Elliot, A. J. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology*, *73*(6), 1284–1295. https://doi.org/10.1037/0022-3514.73.6.1284

Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology*,

94(3), 638–645. https://doi.org/10.1037/0022-0663.94.3.638

Hecker, B. L. (2017). Four decades of open science. *Nature Physics*, *13*(6), 523–525. https://doi.org/10.1038/nphys4160

Henry, F., Dua, E., Kobayashi, A., James, C., Li, P., Ramos, H., & Smith, M. S. (2017). Race, racialization and Indigeneity in Canadian universities. *Race Ethnicity and Education*, *20*(3), 300–314. https://doi.org/10.1080/13613324.2016.1260226

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, *25*(4), 378–405. https://doi.org/10.1006/ceps.1999.1026

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hurtado, S., & Ruiz Alvarado, A. (2012). *The climate for underrepresented groups and diversity on campus* (tech. rep.). HERI Research Brief.

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, *19*(2), 141–184. https://doi.org/10.1007/s10648-006-9012-5

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, *75*(1), 63–82. https://doi.org/10.3102/00346543075001063

Li, K. (2019). MOOC learners' demographics, self-regulated learning strategy, perceived learning and satisfaction: A structural equation modeling approach. *Computers & Education*, *132*, 16–30. https://doi.org/10.1016/j.compedu.2019.01.003

Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology*, *33*(4), 486–512. https://doi.org/10.1016/j.cedpsych.2007.08.001

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, *25*(8), 1608–1618. https://doi.org/10.1177/0956797614535810

McGee, E. O. (2020). Interrogating structural racism in STEM higher education. *Educational Researcher*, *49*(9), 633–644. https://doi.org/10.3102/0013189X20972718

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, *57*(1), 487–503. https://doi.org/10.1146/annurev.psych.56.091103.070258

Merkle, E. C., & Rosseel, Y. (2015). Blavaan: Bayesian structural equation models via parameter expansion. *ArXiv Preprint ArXiv:1511.05604*.

Mouratidis, A., Michou, A., Demircioğlu, A. N., & Sayil, M. (2018). Different goals, different pathways to success: Performance-approach goals as direct and mastery-approach goals as indirect predictors of grades in mathematics. *Learning and Individual Differences*, *61*, 127–135. https://doi.org/10.1016/j.lindif.2017.11.017

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology*, *34*(4), 306–318. https://doi.org/10.1016/j.cedpsych.2009.06.005

Museus, S. D., & Liverman, D. (2010). High-performing institutions and their implications for studying underrepresented minority students in STEM. *New Directions for Institutional Research*, *2010*(148), 17–27. https://doi.org/10.1002/ir.358

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (tech. rep.). ERIC. https://eric.ed.gov/?id=ED338122

Plucker, J. A., & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist*, *56*(2), 90–100. https://doi.org/10.1080/00461520.2021.1895796

Psaltou-Joycey, A. P. A. (2008). Cross-cultural differences in the use of learning strategies by students of greek as a second language. *Journal of Multilingual and Multicultural Development*, *29*(4), 310–324. https://doi.org/10.1080/01434630802147817

Purdie, N., Hattie, J., & Douglas, G. (1996). Student conceptions of learning and their use of self-regulated learning strategies: A cross-cultural comparison. *Journal of Educational Psychology*, *88*(1), 87–100. https://doi.org/10.1037/0022-0663.88.1.87

Ranellucci, J., Muis, K. R., Duffy, M., Wang, X., Sampasivam, L., & Franco, G. M. (2013). To master or

perform? exploring relations between achievement goals and conceptual change learning. *British Journal of Educational Psychology*, *83*(3), 431–451.

Revelle, W. (2018). psych: Procedures for psychological, psychometric, and personality research. *R Package Version*, *1*(10).

Rivers, M. L., & Dunlosky, J. (2020). Are test-expectancy effects better explained by changes in encoding strategies or differential test experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*.

Seaman, C. S., & Weber, R. (2015). Undisclosed flexibility in computing and reporting structural equation models in communication science. *Communication Methods and Measures*, *9*(4), 208–232. https://doi.org/10.1080/19312458.2015.1096329

Senko, C., Hama, H., & Belmonte, K. (2013). Achievement goals, study strategies, and achievement: A test of the "learning agenda" framework. *Learning and Individual Differences*, *24*, 1–10. https://doi.org/10.1016/j.lindif.2012.11.003

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, *46*(1), 26–47. https://doi.org/10.1080/00461520.2011.538646

Senko, C., & Miles, K. M. (2008). Pursuing their own learning agenda: How mastery-oriented students jeopardize their class performance. *Contemporary Educational Psychology*, *33*(4), 561–583. https://doi.org/10.1016/j.cedpsych.2007.12.001

Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educational and Psychological Measurement*, *79*(2), 310–334. https://doi.org/10.1177/0013164418783530

Tabak, F., Nguyen, N., Basuray, T., & Darrow, W. (2009). Exploring the impact of personality on performance: How time-on-task moderates the mediation by self-efficacy. *Personality and Individual Differences*, *47*(8), 823–828. https://doi.org/10.1016/j.paid.2009.06.027

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

van der Zee, T., & Reich, J. (2018). Open education science. *AERA Open*, *4*(3), 2332858418787466. https://doi.org/10.1177/2332858418787466

Vrugt, A., & Oort, F. J. (2008). Metacognition, achievement goals, study strategies and academic achievement: Pathways to achievement. *Metacognition and Learning*, *3*(2), 123–146. https://doi.org/10.1007/s11409-008-9022-4

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, *51*(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

Yan, V. X., & Schuetze, B. A. (2023). What is meant by "growth mindset"? Current theory, measurement practices, and empirical results leave much open to interpretation: Commentary on Macnamara and Burgoyne (2023) and Burnette et al. (2023). *Psychological Bulletin*, (3-4), 206–219. https://doi.org/10.1037/bul0000370