

RESEARCH

Open Access



Assessing the predictive validity of expectancy theory for academic performance

Vahe Permzadian^{1*} and Ting Shen¹

Abstract

Background Despite expectancy theory's widespread appeal and influence as a framework for motivation in organizational and educational settings, studies that have examined the theory's validity for performance-based outcomes, particularly with academic performance as the criterion, have been characterized by inconsistent results. Given numerous methodological concerns associated with past studies (e.g., prevalence of between-person rather than within-person design), we examined the predictive validity of expectancy theory for academic performance using methods that were consistent with the theory's original conceptualization. Additionally, we assessed the validity of the theory for students' study effort.

Methods The final sample included 123 undergraduate students who reported their final grades in four courses. Study effort and other variables were measured with self-report surveys. Because course grades were nested within each person, multilevel modeling was used to test study hypotheses.

Results Both the valence model and the force model predicted a student's current study effort, but contrary to expectations, neither model predicted a student's final course grades. In contrast, both valence for academic success and the simplified force model (based only on valence and expectancy) predicted current study effort, final course grades, and explained incremental variance beyond cognitive ability. Furthermore, the predictive validity of this force model was relatively stable across the 11 weeks of the study.

Conclusions Based on methods congruent with expectancy theory's original framework, we find that the force model does not predict academic performance. An alternative version of the model, however, predicts course grades and has incremental validity over cognitive ability. Our results have several significant theoretical and practical implications.

Keywords Expectancy theory, Motivation, Validity, Academic performance

Expectancy theory [1] has been one of the most influential theories of work motivation in the organizational psychology literature [2, 3]. Influenced by Tolman's [4, 5] claims that choices are made to maximize pleasure and minimize pain, the central tenet of expectancy theory is that our motivation is based on a conscious and rational decision-making process that is intended to maximize the pleasure that can be derived from our choices [1]. The specific process that leads to engaging in any one behavior is determined by evaluating each of the various behavioral alternatives (i.e., choices) concerning the theory's three core constructs of valence, instrumentality,

*Correspondence:

Vahe Permzadian
vpermzadian@mst.edu

¹ Department of Psychological Science, Missouri University of Science and Technology, Rolla, MO 65409, USA



and expectancy. Expectancy theory assumes that individuals are rational actors and will thereby choose to engage in the behavioral alternative that is most motivating.

Expectancy refers to the probability of achieving a desired goal given a reasonable effort. Thus, it can be considered the relationship between effort and goal attainment (e.g., a student believes that studying 5 h for an upcoming exam will lead to an A grade; this reflects a high level of expectancy). The desired goal is typically a specific level of performance in work or achievement settings (e.g., an A grade in a course, an exceptional performance review). Instrumentality refers to the extent that achieving a desired goal is essential before the individual can experience likely outcomes. Thus, it can be considered the relationship between goal achievement and experiencing likely desired outcomes (e.g., an employee does not believe that receiving a favorable performance review is required for a promotion; this reflects a low level of instrumentality). Valence refers to how much an individual values goal attainment based on the anticipated level of satisfaction associated with its likely outcomes (e.g., a business administration student is excited about joining the Business and Economics Club at Muhlenberg College because they believe membership will enhance their resume and the prospects of a future internship: this reflects a high level of valence).

According to Vroom’s [1] original conceptualization of expectancy theory, the motivation to act is determined by valence, instrumentality, and expectancy as mathematically represented by the two component models of valence and force. The valence model, which is based on the constructs of valence and instrumentality, determines the degree to which an individual desires to reach a primary goal (i.e., first level) as a function of the aggregate valences of resulting secondary outcomes (i.e., second level) and the instrumentality associated with these secondary outcomes. As such, a first-level outcome reflects the successful attainment of a current goal, whereas second-level outcomes are the perceived consequences (whether positive or negative) that will likely ensue from goal accomplishment. The computational formula for the valence model is as follows:

$$V_j = f \sum_{k=1}^n (V_k I_{jk}) \tag{1}$$

where,

V_j = Valence (first-level outcome): the value of reaching goal j

I_{jk} = Instrumentality: the extent that reaching goal j is necessary for the attainment of second-level outcome k

V_k = Valence (second-level outcome): the value of second-level outcome k

Table 1 Overview of expectancy theory models

Model	Model components	
Valence	Valence of second-level outcome (V_k)	Instrumentality (I_{jk})
Force	Valence of first-level outcome (V_j)	Expectancy (E_{ij})

Valence model determines the value of achieving a first-level outcome (V_j); force model determines the level of motivation associated with engaging in a behavior (F_i). For brevity, computational formulas are not included

n = the perceived number of second-level outcomes

For example, the value that an undergraduate student might place on attending all classes (i.e., the valence of reaching the goal) is determined by how much the student might desire a high cumulative GPA and graduating in a relatively short period (i.e., the valence of anticipated second-level outcomes) and their perception of how essential attending all classes is to the attainment of good grades and graduating quickly (i.e., the extent that reaching the goal is necessary for the attainment of these second-level outcomes).

The force model, which is based on the valence model and the construct of expectancy, determines the level of motivation (referred to as “force” by Vroom [1]) that is associated with engaging in any one behavior as a function of the valence of reaching the goal (based on V_j from the valence model) and the perceived likelihood of reaching the goal given a reasonable effort. The computational formula for the force model is as follows:

$$F_i = \sum_{j=1}^n (E_{ij} V_j) \tag{2}$$

where,

F_i = Force: the level of motivation that is associated with engaging in behavioral alternative i

E_{ij} = Expectancy: the likelihood that choosing to engage in behavioral alternative i (i.e., a certain level of effort toward an act) will lead to the attainment of goal j

V_j = Valence (first-level outcome): the value of reaching goal j

Returning to the previous example, the level of motivation that is associated with attending all classes is determined by the student’s perception of the probability of being able to successfully attend all classes assuming that they are putting forth a reasonable level of effort and how much they desire successfully attending all classes (i.e., the valence of reaching goal j , which is based on the valence model). Table 1 summarizes the valence and force models.

Although based on the force model, the level of motivation for any one behavior is represented by a multiplicative function (i.e., expectancy \times valence), research examining the type of information processing that

individuals employ (e.g., [6–9]) has found that an additive processing model is more accurate for the majority of participants as compared to a multiplicative processing model [2, 3]. Thus, there are individual differences regarding how these components are combined. However, most assessments of expectancy theory have remained consistent with Vroom [1] and thereby measured the force model with a multiplicative function.

There is support for expectancy theory's major proposition that motivation to engage in a behavior is determined by an individual's perceptions of valence, instrumentality, and expectancy (e.g., [10–14]). Furthermore, these findings have been observed in both the laboratory (e.g., [7]) and field settings (e.g., [15]), across various cultures (e.g., [16]), and different populations (e.g., [17–19]). There are several important observations from this vast body of literature. Although there has been moderate support for expectancy theory's main tenets, the literature is also characterized by a high level of variability in results, as demonstrated by studies that have reported non-significant effects as well as small-to-large positive effects [2, 13, 20–24]. It is essential to highlight that a vast body of literature in educational psychology has examined a theory with similarities to expectancy theory. Expectancy-value theory [25–27], an extension of Atkinson's [28, 29] expectancy-value model, was proposed specifically to explain the motivation of students. There has been more consistent support for the validity of expectancy-value theory for performance in higher education [30–33].

There are several reasons for the highly variable results in studies that have assessed the validity of expectancy theory. First, study findings have largely differed according to the outcome variables that have been used in assessments of the theory [3, 13]. The common criteria have included performance (e.g., grades in school, job performance), effort (e.g., the amount of time spent on a task), intention (e.g., desire to apply for a job), preference (e.g., how attractive an occupation is as compared to others), and choice (e.g., choosing to leave an organization). Based on meta-analytic results by Van Eerde and Thierry [13], expectancy theory components have a stronger relationship with attitudinal criterion variables (e.g., intention, preference) rather than behavioral (e.g., performance, effort, choice). The results with performance as the outcome have been particularly poor [2, 3]. Based on the theory of planned behavior [34, 35], this variability in results is perhaps not surprising when considering that the relationship between a focal predictor and an outcome would be expected to be stronger for attitudinal outcomes as compared to behavioral outcomes.

Second, results have largely differed based on the study design that has been used to assess the theory's predictive

power. Some studies have used a within-person design (e.g., [36]), whereas others have used a between-person design (e.g., [17]). Based on Vroom [1], however, the choice to engage in any one behavioral alternative based on its motivational force is relative to the perceived motivational force of all other behavioral alternatives. As such, Vroom viewed the relationship between expectancy theory variables and criteria as properly involving a within-person analysis [13, 23]. Hence, the theory was never intended to predict differences in motivation between individuals. However, most studies have used a between-person design. Consistent with Vroom's view, studies using a within-person design have reported higher correlations than those using a between-person design with effort or preference as criteria [13]. It might be argued that, to some extent, this greater reliance on between-person designs is an explanation for the observed weak relationship between expectancy theory components and performance.

Third, study findings have largely differed based on highly variable operationalizations of expectancy theory variables and, in particular, the construct of valence [2, 13]. For example, valence has been conceptualized as reflecting attractiveness, importance, and desirability. Further, there have been problems with measuring the valence model. According to the computational formula for the valence model (see Eq. 1), the valence of successfully attaining a current goal is based on the aggregate valences of anticipated secondary outcomes and the instrumentality associated with these secondary outcomes. As discussed by Vroom [1], an important implication of this equation is that individuals can vary concerning (1) the specific second-level outcomes that they would expect to result from the attainment of the first-level outcome, and (2) the number of expected second-level outcomes. However, rarely have studies in this domain provided participants with the opportunity to generate their second-level outcomes. Instead, participants have been expected to choose from a list of outcomes shared by the researcher, which increases the chances that some of the outcomes rated would not have been as relevant as others to participants [13]. To the best of our knowledge, only Matsui and Ikeda [37] gave participants the chance to produce their second-level outcomes and found that the validity of expectancy theory was higher with participant-generated outcomes as compared to researcher-generated outcomes.

Although the literature has been characterized by inconsistent results, expectancy theory's applicability to employees in organizations has arguably been better demonstrated than other populations. Indeed, the theory's validity for the college student population has been lower, particularly with academic performance as a

criterion. Hence, the primary focus of our study is on the predictive validity of expectancy theory for performance in higher education.

Literature review

The review of the expectancy theory literature (with performance as the criterion) has revealed inconsistent findings, with only a few studies reporting a significant relationship between expectancy theory components and academic performance (e.g., [19]). Harrell et al. [38] examined the relationship between undergraduate students' motivation (based on the force model) and average grades over two semesters in intermediate accounting. For students with low and medium expectancy, there was a significant relationship with grades. However, non-significant results were observed for students with high expectancy. These results are not consistent with expectancy theory given that low levels of expectancy predicted grades rather than high. Tyagi [19] explored the relationship between undergraduate students' motivation (based on the force model) and grades in a marketing course, with perceived control (i.e., perceived ease of performing a behavior) as a moderator variable. Motivation was found to predict grades but only for students who had high perceived control.

The next group of studies differed as they also measured cognitive ability, which is necessary if attempting to demonstrate the incremental validity of motivation as assessed by expectancy theory. Youssef [39] examined whether undergraduate students' motivation (based on the theory's three components) would predict grades in foreign language courses. Cognitive ability was measured by a combination of high school GPA and SAT scores. Whereas ability was found to be a significant predictor of grades, motivation was not associated with grades. Malloch and Michael [40] investigated the relationship between undergraduate students' motivation and end-of-term grades. Cognitive ability was measured by composite ACT or SAT scores. Although the study found that both motivation and cognitive ability were significant predictors of grades, with ability accounting for a greater proportion of the variance, it is difficult to consider it as an appropriate test and application of expectancy theory given that several measures of study variables were not consistent with Vroom [1]. For example, expectancy was operationalized as a student's anticipated end-of-term GPA rather than the likelihood of goal success. Pringle [41] explored whether undergraduate students' motivation (based on the force model) would predict grades in an organizational behavior class. Cognitive ability was assessed with composite SAT scores. Similar to Youssef [39], ability was found to be a significant predictor of grades, but motivation was not associated with grades.

In summary, reviewing the literature that has adopted expectancy theory and examined its predictive validity for academic performance shows that only two studies have found results in support of and consistent with expectancy theory [19, 40]. However, given that the effect of motivation on grades was moderated by another variable in Tyagi [19] and the observation that Malloch and Michael's [40] measurement of expectancy theory variables was not in line with Vroom's [1] operationalizations, we can conclude that the level of support for the theory is weak.

We believe that at least some of the apparent low validity of the theory in predicting grades stems from the various issues we detailed for the inconsistent results that have been observed (e.g., operationalizations of expectancy theory constructs that are not in line with Vroom [1]). Valence, in particular, has been problematic as many studies have used scales with only positive anchors in its measurement rather than scales with both positive and negative anchors [13]. As Donovan [2] noted:

Despite the large number of studies that have set out to assess the validity of this theory, we are still left with very few studies that actually test this theory in an appropriate manner. As such, it appears that we have yet to convincingly answer some of the basic questions concerning the accuracy of this model, suggesting that this research area would benefit greatly from additional, quality research assessing the validity of this model. (p. 59)

Present study

To address the inconsistent findings in the literature and the low predictive validity of the theory for performance in higher education, the main purpose of this study is to examine expectancy theory's validity for academic performance based on Vroom's [1] original framework. As a secondary goal, we also assess the theory's validity for students' study effort. In addition to assessing the component models of valence and force, we also examine the construct of valence and a variant of the force model that does not include instrumentality, and thus is based only on valence and expectancy (described further in subsequent sections). To properly assess expectancy theory's validity, our study (1) measures the theory's core variables as specified by Vroom, (2) provides participants with the opportunity to produce the second-level outcomes that they deem appropriate, (3) assesses hypotheses using a within-person design, (4) includes specific course grades as criteria rather than the mean grade across courses, and (5) models the predictive validity of the theory over time. Additionally, our study examines the incremental validity of expectancy theory beyond cognitive ability, which

has not been demonstrated in the literature.¹ To the best of our knowledge, this is the first study to (1) measure the valence model precisely based on Vroom by allowing participants to list any second-level outcome they perceived as possible (i.e., both positive and negative outcomes were possible) and allowing for individual differences in the possible number of second-level outcomes that participants might generate, (2) model change across time in the validity of the force model using a within-person design, and (3) use specific course grades as the outcome rather than an average term GPA, which allows for greater precision in analysis.

Theoretical framework and hypotheses development

According to expectancy theory [1], the motivation to choose any one behavior is partly determined by an individual's desire for reaching a current goal (based on the valence model) and partly determined by the perceived expectations of goal success (based on expectancy in the force model). Hence, evaluations of the theory can involve the full model (i.e., testing both the valence model and the force model), assessing components of the theory (e.g., only testing the valence model), or assessing specific expectancy theory variables [13].

Regardless of the specific theory of motivation, typical definitions of motivation describe the construct as being reflected by the direction, intensity, and duration of behavior (e.g., [42–44]). Intensity and magnitude of behavior are synonymous with and reflective of an individual's effort. Therefore, the most direct assessment of expectancy theory involves examining if it predicts the level of effort. Expectancy theory's models have been found to predict effort in the sparse research with students in higher education (e.g., [45]). Hence the following hypotheses are proposed:

Hypothesis 1a: The valence model will explain a student's current level of study effort.

Hypothesis 1b: The force model will explain a student's current level of study effort.

One of the major assumptions of expectancy theory is that individuals are rational actors who make optimal decisions by weighing each behavioral alternative's expectancy, valence, and instrumentality [3]. This critical assumption rests on a related assumption that individuals are fully aware of all behavioral alternatives (i.e., first-level

outcomes) and their perceived consequences (i.e., second-level outcomes). Unfortunately, these assumptions are unlikely to be met and to apply in all instances [46]. Depending on the circumstances and their cognitive resources, individuals are also likely to use variations of expectancy theory's models that are simpler and do not entail comprehensive processing [47]. Indeed, the valence and force models are associated with smaller effect sizes as compared to individual expectancy theory constructs and thereby using specific constructs has been recommended [13]. Accordingly, we also examined the construct of valence and a modified force model that does not include instrumentality and relies on valence and expectancy. This simplified version of expectancy theory is arguably the most elementary version of the model as it retains its most fundamental components (i.e., motivation = expectancy × valence). Thus, the following hypotheses are proposed:

Hypothesis 2a: A student's perceived valence for academic success will be positively associated with their current level of study effort.

Hypothesis 2b: The simplified force model will explain a student's current level of study effort.

Expectancy theory was proposed to predict choice, intention, and effort rather than performance [2, 22]. Regardless, any framework that assesses motivation would be expected to also predict performance in academic and organizational settings based on the vast body of literature that has theoretically proposed and empirically supported motivation and ability as the two primary determinants of performance (e.g., [21, 43, 48–58]). Decades of research have established that the mechanism through which motivation affects behavior in general and performance in particular is by promoting the attainment of performance-related goals [59–61]. Hence the following hypotheses are proposed:

Hypothesis 3a: The valence model will explain a student's final course grades.

Hypothesis 3b: The force model will explain a student's final course grades.

A substantial amount of literature indicates that general cognitive ability and motivation are two of the primary determinants of academic success (e.g., [62–67]). Furthermore, there is extensive research evidence that motivation and related constructs (e.g., achievement goals) have incremental validity beyond cognitive ability in predicting academic achievement (e.g., [67–75]). Thus, the following hypotheses are proposed:

¹ Although Malloch and Michael [40] found that motivation and cognitive ability were significant predictors of grades, several measures of study variables were not consistent with Vroom [1] and thus difficult to classify it as an appropriate test of the theory's validity.

Hypothesis 4a: The valence model will explain incremental variance in a student's final course grades beyond cognitive ability.

Hypothesis 4b: The force model will explain incremental variance in a student's final course grades beyond cognitive ability.

As discussed in the rationale that preceded Hypothesis 2a and 2b, depending on the situation, individuals may use variations of expectancy theory's component models that are simpler and thereby devoid of exhaustive information processing [47]. Consistent with this argument, using the theory's specific constructs has been recommended [13]. Further, based on our arguments that preceded Hypothesis 3a and 3b, a model assessing motivation would be expected to predict performance based on the volume of studies that have found motivation and ability as the primary determinants of performance (e.g., [43, 51, 54, 62, 64, 66]). Thus, the following hypotheses are proposed:

Hypothesis 5a: A student's perceived valence for academic success will be positively associated with their final course grades.

Hypothesis 5b: The simplified force model will explain a student's final course grades.

The empirical evidence from expectancy theory studies in higher education that have assessed its incremental validity beyond cognitive ability is poor (e.g., [39, 41]). Given our arguments for also assessing a simpler version of the theory that includes perceived valence, it would be appropriate to refer to research that has examined variables similar to valence. In the expectancy-value theory literature [25–27], extensive work has been conducted on attainment value (i.e., the importance of doing well on a certain task), which is nearly identical to the operationalization of valence as reflecting the importance of a goal. Research on the effect of attainment value has found that it is associated with academic achievement and explains incremental variance beyond cognitive ability (e.g., [76]). Thus, the following hypotheses are proposed:

Hypothesis 6a: A student's perceived valence for academic success will explain incremental variance in their final course grades beyond cognitive ability.

Hypothesis 6b: The simplified force model will explain incremental variance in a student's final course grades beyond cognitive ability.

Method

Participants

The initial study participants were 302 undergraduate students at a large Northeastern US university. When contacted at the end of the term, 123 students reported their final course grades in four courses (response rate of 40.7%) and were included in the final sample. The mean age of these participants was 18.5 years. Approximately 63% identified as female and 37% identified as male. Participants' academic majors consisted of social sciences (21.9%), humanities (9.8%), business (30.1%), biological sciences (20.3%), mathematics/computer science (4.1%), and the remaining 13.8% were undecided. In terms of class standing, 63.4% were first-year students, 23.6% were second-year students, 9.8% third-year students, and 3.2% fourth-year students. Regarding ethnicity and racial composition, approximately 59.3% of participants identified as White, 12.2% as Black, 14.6% as Hispanic, 9.0% as Asian, and 4.9% identified as "other."

We conducted a series of tests to compare participants who were included in the final sample and those who were only part of the initial sample. Based on these analyses, the two groups did not differ significantly in age: $t(300) = -.06, p = .95$; mean class standing: $t(300) = .21, p = .83$; or cognitive ability: $t(300) = -1.14, p = .25$. There was, however, a difference in gender as the final sample included more participants who identified as female (62.6% versus 46.3%, $p < .01$).

Procedure

Study participants completed paper-based surveys in a laboratory at the university. After the study was advertised by the research pool (a few weeks into the term), students could participate until the final weeks of the term; thus, data collection lasted 11 weeks. Except for a few weeks, there was a consistent number of participants (approximately 14 per week).

After giving informed consent to participate in the study, participants were made aware that the purpose of the study was to understand the study activities and behaviors of students and the factors that influence how individuals make study-related decisions. Participants were also informed that they would be contacted via email at the end of the term by the principal investigator and instructed to share their final grades in four 3-credit courses. Participants then completed the initial study survey that collected biographical information and included a portion to report SAT scores. Next, participants completed the main study survey that measured valence and force model variables as well as study effort and academic locus of control (as a control variable). This survey also assessed valence, which was needed for the

modified force model. Participants had 45 min to complete the different study surveys and were granted 1 h of research credit for taking part in the study. At the end of the term, the principal investigator contacted study participants to request their course grades. To provide an incentive to respond, all participants were entered into a raffle for three \$25 Amazon gift cards.

Measures

Valence (second-level outcomes)

Based on Vroom [1], the valence model determines how much an individual desires reaching a first-level outcome (i.e., a current goal) as a function of the aggregate valences of second-level outcomes (i.e., perceived consequences of a first-level outcome) and the instrumentality associated with the second-level outcomes (see Eq. 1). The valence of a second-level outcome is defined as an individual's anticipated level of satisfaction associated with a particular outcome [1]. Given that this study assesses the motivation of college students for academic success, the first-level outcome was an "A" grade in each of four different courses and thereby the second-level outcomes were participants' perceived consequences of receiving this grade in a respective class. Specifically, to measure the valence of second-level outcomes, participants were first given the opportunity to imagine receiving an "A" grade at the end of the term in each of the four classes and then listing the outcomes (up to seven) that they anticipate from this grade for every class (whether positive or negative outcomes). Next, participants were instructed to imagine experiencing each of the different outcomes at the end of the term for a respective course and rate their anticipated level of satisfaction on a scale ranging from +10 (*very satisfied*) to -10 (*very dissatisfied*). Thus, consistent with Vroom, we measured valence by (1) operationalizing it as reflecting the anticipated satisfaction with an outcome, (2) providing participants with the chance to generate second-level outcomes that they deemed as most relevant, and (3) using a scale with both positive and negative anchors to satisfy Vroom's assumption that the construct can also assume negative values.

Instrumentality

Instrumentality is defined as the extent that reaching a first-level outcome is necessary for an individual to attain anticipated second-level outcomes [1]. Thus, it can be considered as the perceived strength of the relationship between a first-level outcome and its associated second-level outcomes. To assess instrumentality, participants had to first refer to the expected outcomes of an "A" grade in a respective class that they had listed (i.e., second-level outcomes for each class). Next, participants rated the extent that receiving an "A" grade in a respective class

affected their chances of experiencing each outcome on a scale ranging from +10 (*highly increases*) to -10 (*highly decreases*). Consistent with Vroom [1], we measured instrumentality by using a scale with both positive and negative anchors to satisfy Vroom's assumption that the variable can assume both positive and negative values.

Expectancy

According to Vroom [1], expectancy refers to an individual's belief regarding the probability that a certain level of effort will lead to the attainment of a first-level outcome (i.e., a desired goal). Thus, it can be considered as the relationship between an action (i.e., effort) and an outcome (i.e., goal success). To assess expectancy, participants indicated the probability that studying hard (i.e., a high level of study effort) would lead to an "A" grade in a respective class on a scale ranging from 0-100%.

Valence

Vroom [1] viewed valence as representing an individual's desire or preference for a first-level outcome (i.e., a current goal). To assess valence, participants rated how much they desired to receive an "A" grade in a respective class on a scale ranging from +10 (*highly desire receiving it*) to -10 (*highly desire NOT receiving it*).

Effort

To measure effort in each course, participants reported the amount of time they were currently allocating to study activities. These activities were defined as "reading, writing, listening to recorded lectures, attending tutoring, or any other study activity not listed that helps your performance in this class." Participants provided an estimate of the total (combined) number of hours per week, ranging from 0-20 h, that they were currently spending on various study activities for a class. If the number of hours exceeded 20, participants were instructed to indicate the specific number.

Academic performance

Academic performance was operationalized as a student's self-reported end-of-term grades. Participants reported their final grades (A-F) in four courses at the end of the academic term. Self-reported grades have been found to have validity and are reliable substitutes when accessing grades from school records is not practical (e.g., [77]).

Cognitive ability

Cognitive ability was operationalized as a student's self-reported SAT score in reading, math, and writing. The SAT composite score (summed total) was then used as a proxy measure of cognitive ability. Standardized aptitude tests such as the SAT and ACT have been found to

be valid measures of general cognitive ability (e.g., [78–81]). Indeed, there is a high degree of empirical overlap between these assessments of scholastic aptitude and general cognitive ability assessments [82]. Self-reported SAT scores have also been found to have validity and are reliable substitutes when accessing scores from school records is not feasible (e.g., [77, 83, 84]).

Control variable

Although only a few studies have examined locus of control as a moderator of expectancy theory constructs and outcomes such as effort and performance (e.g., [85]), we measured locus of control because it has been linked with both expectancy (e.g., [85, 86]) and instrumentality (e.g., [87]) as well as effort (e.g., [88]) and performance (e.g., [88, 89]) and thus conceptually can operate as a moderating variable [90, 91]. Locus of control was measured with the 28-item Academic Locus of Control scale (ALOC) [92] that uses a true-false response format and was specifically designed to assess college students' beliefs about perceived control over academic outcomes. Higher ALOC scores represent a higher external locus of control (i.e., the belief that outcomes are beyond an individual's control), whereas lower scores represent a higher internal locus of control (i.e., the belief that outcomes are within an individual's control). A sample item was "some people have a knack for writing, while others will never write well no matter how hard they try." Cronbach's alpha was .72 in our study.

Analytic strategy

Valence model

Based on the computational formula for the valence model (see Eq. 1), we calculated each participant's V_j (i.e., the valence of reaching goal j [a first-level outcome]) by (1) multiplying the valence of each second-level outcome (V_k) by the instrumentality associated with each second-level outcome (I_{jk}), (2) summing across the number of second-level outcomes (n), and (3) dividing by the number of second-level outcomes (n) to determine the mean. The third step (calculating the average V_j score) was necessary given that the number of second-level outcomes generated varied across participants and thus omitting this step could have led to higher V_j scores for participants who, by chance, had listed more outcomes.

Given research evidence that many participants integrate expectancy theory components additively (e.g., [8]) rather than multiplicatively, a valence model based on an additive integration of variables was also examined. We had no a priori hypothesis regarding the type of information processing as it was not the study's focus. Thus, we conducted this additional analysis to compare the validity

of the valence model with multiplicative processing against the model with additive processing.

Based on the computational formula for the valence model, the additive valence model was calculated using the same steps that were used for the multiplicative valence model. The only difference was that the first step involved adding the valence of each second-level outcome (V_k) to the instrumentality associated with each second-level outcome (I_{jk}).

Force model

Based on the computational formula for the force model (see Eq. 2), we calculated each participant's F_i (i.e., level of motivation that is associated with engaging in behavioral alternative i) by multiplying their perceived expectancy (E_{ij}) by the valence of reaching goal j (V_j), which was derived from the valence model.

We also examined a force model based on an additive integration of variables. As before, we had no a priori hypothesis regarding processing type and thereby conducted this additional analysis to compare the validity of the force model with multiplicative processing against the model with additive processing.

Based on the force model's computational formula, we calculated each participant's additive F_i by (1) converting the raw expectancy score (E_{ij}) and the raw valence score from the additive valence model (V_j) to a z score and (2) adding these two z scores. The first step (converting to z scores) was necessary given that these two variables were on different scales and adding their raw scores would not be correct.

Simplified force model

We calculated each participant's F_i based on the modified force model by multiplying their perceived expectancy (E_{ij}) by the direct measure of valence. We also examined a simplified force model based on an additive integration of variables. As before, this additional analysis was conducted to compare the validity of the simplified force model with multiplicative processing against the same model with additive processing. We calculated each participant's additive F_i of this modified force model by using the same steps that were used to calculate the additive force model. Specifically, we (1) converted the raw expectancy score (E_{ij}) and the raw valence score to a z score and then (2) added these two z scores.

Predictive validity

To model the predictive validity of expectancy theory, we examined how the magnitude of the force model's unstandardized regression coefficients (both multiplicative and additive) changed over the length of our study, which lasted for 11 weeks (i.e., the length of time

between the first and last participant who completed the study surveys). This analysis involved (1) dividing participants (based on their date [i.e., specific week] of study participation) into 11 groups by weeks (e.g., Week 1 includes participants who took part in the first 7 days of the study), (2) using multilevel modeling (MLM) to assess the strength of the relationship between the force model and final course grades for each of these 11 weeks, and (3) plotting the force model's regression coefficients to assess change across weeks.

Data structure and analyses

Given the hierarchical structure of our data, with course-level data (i.e., Level 1) nested within persons at Level 2, MLM [93] was used to test the study hypotheses. All Level-1 predictors (i.e., valence model, force model, valence, and the simplified force model) were person-mean centered. The following steps were used in building a multilevel model to test all hypotheses. In Step 1, we assessed a null model with only a criterion variable (e.g., study effort). The proportion of variance within- and between-person was estimated by the intraclass correlation coefficient (ICC). Based on the ICCs of .56 and .26 (see Tables 3, 4 and 5), the use of MLM in the current study was warranted as 44-74% of the variance in study variables was at the within-person level. In Step 2, we added any Level-1 covariates (i.e., ALOC) to the model. As recommended by Nezlek [94] and Bernerth and Aguinis [95], predictors that were not significant were removed from the model before any new predictors were added in subsequent models.² In Step 3, we estimated a random intercept and fixed slope model by adding Level-1 predictors. In Step 4, incremental validity was assessed with the addition of cognitive ability to the preceding model. Pseudo R^2 (within-person level) was reported for models where appropriate [96]. We used SPSS 21 to conduct all MLM analyses and descriptive statistics.

Results

Multiplicative versus additive models

We had no a priori hypotheses regarding the validity of expectancy theory models with a multiplicative versus an additive integration of variables. Based on our results, there was no difference in the validity of a multiplicative valence model and an additive valence model (i.e., when the valence model was the focal predictor in Hypotheses 1a, 3a, and 4a) as support or lack thereof for these three hypotheses was the same regardless of whether the model used a multiplicative or an additive integration.

For example, Hypothesis 1a was supported using both the multiplicative and the additive valence model. Alternatively, there was no support for Hypothesis 3a using either the multiplicative or the additive valence model. There was also no difference in the validity of a multiplicative force model and an additive force model (i.e., when the force model was the focal predictor in Hypotheses 1b-6b) as support or lack thereof for these six hypotheses was the same regardless of how the theory's variables were integrated.

Because of this lack of difference between the validity of multiplicative and additive models, we included the unstandardized regression coefficients and corresponding p values of multiplicative models in our discussion of study results as well as results in Tables 3, 4 and 5 to remain consistent with Vroom's [1] original expectancy theory model.

Data screening

Quantile-quantile (Q-Q) plots were used to check the normality of residuals assumption of multilevel modeling. Based on the Q-Q plots, none of the inspected variables exhibited obvious or severe violations of this assumption. Univariate outliers were checked by identifying cases with a z-score of $+/-3.29$ [97]. Two participants responded to several study variables in a manner that can be considered outliers. Accordingly, these responses were inspected based on recommendations by Aguinis et al. [98]. The two cases were determined to likely represent accurate points of data that happen to stand apart from other data points. Therefore, we decided against removing these cases.

Descriptive statistics and hypotheses testing

Table 2 presents the means, standard deviations, and correlations among study variables. Hypothesis 1a predicted the valence model correlates with study effort. In support of this hypothesis, the valence model ($\gamma_{20}=0.028$, $p<.001$, pseudo $R^2=.028$) was positively associated with effort (see Table 3, Model 3a). Hypothesis 1b predicted the force model correlates with study effort (Hypothesis 1b). This hypothesis was also supported as the force model ($\gamma_{20}=0.039$, $p<.001$, pseudo $R^2=.063$) was positively associated with effort (see Table 3, Model 3b). Hypothesis 2a predicted perceived valence positively correlates with study effort. In support of this hypothesis, valence ($\gamma_{20}=0.237$, $p<.01$, pseudo $R^2=.025$) was positively associated with effort (see Table 3, Model 3c). Hypothesis 2b predicted the simplified force model correlates with study effort. The simplified force model ($\gamma_{20}=0.303$, $p<.001$, pseudo $R^2=.058$) was positively associated with effort (see Table 3, Model 3d), thus supporting the hypothesis.

² The only exception was when assessing incremental validity in Model 4.

Table 2 Means, standard deviations, and correlations among study variables

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8
1. Valence model	70.34	20.72		.76**	.19**	.15**			.17**	.06
2. Force model	60.80	22.35	.92**		.27**	.58**			.26**	.10*
3. Valence	8.86	1.93	.33**	.36**		.73**			.17**	.18**
4. Simplified force model	7.66	2.09	.42**	.63**	.81**				.25**	.13**
5. Locus of control	10.89	4.38	-.24**	-.26**	-.22*	-.21*				
6. Cognitive ability	1639.48	167.02	-.16	-.13	.13	.05	.08			
7. Effort	5.30	3.30	.01	.00	.11	.05	-.04	-.20*		-.02
8. Academic performance	3.26	0.54	.06	.07	.24**	.19*	-.22*	.30**	-.03	

Note. *N* = 123; number of observations = 492. Correlations above the diagonal are within-person correlations. Correlations below the diagonal are between-person correlations

* *p* < .05. ** *p* < .01

Table 3 Multilevel modeling estimates of the effect of expectancy theory predictors on effort

Variable	Null (Model 1)	Covariate (Model 2)	RIFSM (Model 3a)	RIFSM (Model 3b)	RIFSM (Model 3c)	RIFSM (Model 3d)
Fixed effects						
Intercept (γ_{00})	5.299 (0.297)***	5.649 (0.802)***	5.299 (0.297)***	5.299 (0.297)***	5.299 (0.297)***	5.299 (0.297)***
ALOC (γ_{10})		-0.032 (0.068)				
Valence model (γ_{20})			0.028 (0.008)***			
Force model (γ_{20})				0.039 (0.008)***		
Valence (γ_{20})					0.237 (0.074)**	
Simplified force model (γ_{20})						0.303 (0.062)***
Variance components						
Within-person (L1) variance (σ^2)	7.064 (0.520)***	7.064 (0.520)***	6.869 (0.506)***	6.618 (0.488)***	6.890 (0.508)***	6.652 (0.490)***
Intercept (L2) variance (τ_{00})	9.103 (1.398)***	9.173 (1.412)***	9.152 (1.397)***	9.215 (1.397)***	9.147 (1.397)***	9.206 (1.397)***
ICC	.563					
R^2			.028	.063	.025	.058

Note. γ = unstandardized coefficient estimate; RIFSM random intercept and fixed slope model; ALOC academic locus of control; L1 Level 1; L2 Level 2. L1 *N* = 492; L2 sample size = 123. For brevity, the analyses for valence model, force model, valence, and simplified force model are presented in Models 3a-3d, respectively. Values in parentheses are standard errors

* *p* < .05. ** *p* < .01. *** *p* < .001

Hypothesis 3a predicted the valence model correlates with course grades, and Hypothesis 4a predicted it explains incremental variance beyond cognitive ability. These hypotheses were not supported as the valence model ($\gamma_{20} = 0.001$, $p = .59$) was not associated with grades and did not account for incremental variance (see Table 4, Model 3a and 4a). Hypothesis 3b predicted the force model correlates with grades, and Hypothesis 4b predicted the model explains incremental variance. These hypotheses were also not supported as the force model ($\gamma_{20} = 0.002$, $p = .25$) was not associated with grades and did not account for incremental variance (see Table 4, Model 3b and 4b).

Hypothesis 5a predicted valence correlates with grades, and Hypothesis 6a predicted it explains incremental variance. In support of these hypotheses,

valence ($\gamma_{20} = 0.056$, $p < .01$, pseudo $R^2 = .022$) was positively associated with grades and accounted for incremental variance (see Table 5, Model 3a and 4a). Hypothesis 5b predicted the simplified force model correlates with grades, and Hypothesis 6b predicted the model explains incremental variance. These hypotheses were also supported as the simplified force model ($\gamma_{20} = 0.035$, $p < .05$, pseudo $R^2 = .011$) was positively associated with course grades and accounted for incremental variance (see Table 5, Model 3b and 4b).

Predictive validity over time

Because the traditional force model did not predict a student’s final course grades (Hypothesis 3b), we modeled the predictive validity of the simplified force model in Fig. 1 (after accounting for the effects of

Table 4 Multilevel modeling estimates of the effect of valence model and force model on academic performance

Variable	Null (Model 1)	Covariate (Model 2)	RIFSM (Model 3a)	RIFSM (Model 3b)	Incremental validity (Model 4a)	Incremental validity (Model 4b)
Fixed effects						
Intercept (γ_{00})	3.264 (0.048)***	3.554 (0.126)***	3.553 (0.126)***	3.555 (0.126)***	1.895 (0.448)***	1.893 (0.447)***
ALOC (γ_{10})		-0.027 (0.011)*	-0.027 (0.011)*	-0.027 (0.011)*	-0.030 (0.010)**	-0.030 (0.010)**
Valence model (γ_{20})			0.001 (0.002)		0.001 (0.002)	
Force model (γ_{20})				0.002 (0.002)		0.002 (0.002)
Cognitive ability (γ_{30})					0.001 (0.000)***	0.001 (0.000)***
Variance components						
Within-person (L1) variance (σ^2)	0.461 (0.035)***	0.461 (0.035)***	0.462 (0.035)***	0.461 (0.035)***	0.463 (0.035)***	0.462 (0.035)***
Intercept (L2) variance (τ_{00})	0.162 (0.038)***	0.151 (0.037)***	0.150 (0.037)***	0.150 (0.037)***	0.121 (0.033)***	0.121 (0.033)***
ICC	.260					

Note. γ = unstandardized coefficient estimate; RIFSM random intercept and fixed slope model; ALOC academic locus of control; L1 Level 1; L2 Level 2. L1 N=492; L2 sample size = 123. For brevity, valence model and force model analyses are presented in Model 3a-4a and 3b-4b, respectively. Values in parentheses are standard errors

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 5 Multilevel modeling estimates of the effect of valence and simplified force model on academic performance

Variable	Null (Model 1)	Covariate (Model 2)	RIFSM (Model 3a)	RIFSM (Model 3b)	Incremental validity (Model 4a)	Incremental validity (Model 4b)
Fixed effects						
Intercept (γ_{00})	3.264 (0.048)***	3.554 (0.126)***	3.557 (0.126)***	3.557 (0.126)***	1.870 (0.448)***	1.885 (0.448)***
ALOC (γ_{10})		-0.027 (0.011)*	-0.027 (0.011)*	-0.027 (0.011)*	-0.030 (0.010)**	-0.030 (0.010)**
Valence (γ_{20})			0.056 (0.019)**		0.058 (0.019)**	
Simplified force model (γ_{20})				0.035 (0.017)*		0.036 (0.017)*
Cognitive ability (γ_{30})					0.001 (0.000)***	0.001 (0.000)***
Variance components						
Within-person (L1) variance (σ^2)	0.461 (0.035)***	0.461 (0.035)***	0.451 (0.034)***	0.456 (0.034)***	0.452 (0.034)***	0.457 (0.035)***
Intercept (L2) variance (τ_{00})	0.162 (0.038)***	0.151 (0.037)***	0.155 (0.037)***	0.152 (0.037)***	0.125 (0.033)***	0.123 (0.033)***
ICC	.260					
R^2			.022	.011	.020	.009

Note. γ = unstandardized coefficient estimate; RIFSM random intercept and fixed slope model; ALOC academic locus of control; L1 Level 1; L2 Level 2. L1 N=492; L2 sample size = 123. For brevity, valence and simplified force model analyses are presented in Model 3a-4a and 3b-4b, respectively. Values in parentheses are standard errors

* $p < .05$. ** $p < .01$. *** $p < .001$

locus of control and cognitive ability) as it was associated with course grades (Hypothesis 5b). Consistent with the rest of the results, the figure was based on a multiplicative integration of variables. As observed in Fig. 1, except for Week 5 and Week 8, the magnitude of the simplified force model’s regression coefficients was mostly stable across the 11 weeks of the study. We should note that Week 8 data is based on only two students and thereby susceptible to the influence of outliers.

Discussion

Despite the prominence of Vroom’s [1] expectancy theory as a framework for motivation in organizational and educational settings, studies that have assessed the theory’s validity have found inconsistent and highly variable results [3, 13, 20, 23]. Indeed, studies in higher education have generally failed to predict academic performance or find support for the incremental validity of motivation (as measured by expectancy theory) over cognitive ability (e.g., [39, 41]). However, because of

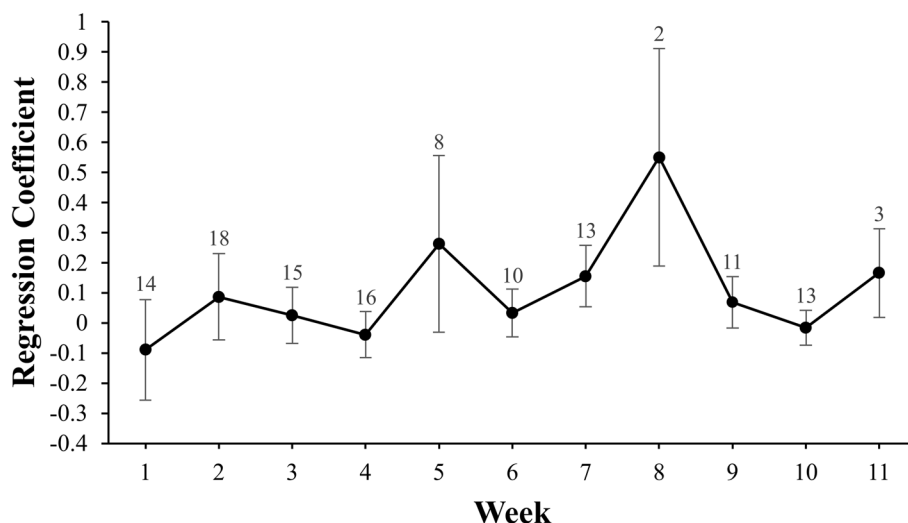


Fig. 1 Predictive validity of simplified force model for academic performance. *Note.* Unstandardized coefficients are based on a person-mean-centered predictor and reflect the validity of the model beyond the effects of locus of control and cognitive ability. Error bars represent 95% confidence intervals. Number of participants per week is displayed above error bars

several methodological concerns associated with studies in this domain (e.g., measurements of the theory’s constructs that are incongruent with Vroom, the prevalence of between-person designs) and in response to calls for more research examining expectancy theory’s validity, our study extended previous research in several ways (e.g., measured the valence model by permitting participants to list any second-level outcome they perceived as relevant) and assessed the theory’s validity.

Consistent with expectancy theory, the valence model and the force model were both associated with a student’s current study effort as Hypotheses 1a and 1b were fully supported. These results are in line with previous research that has found instrumentality (i.e., a component of the valence model) and the force model are positively associated with anticipated effort [7, 45]. However, our results extend these previous studies by finding that expectancy theory also predicts the present level of study effort that is exerted by students.

Our results are also supportive of the most elementary version of expectancy theory as valence and the simplified force model (based only on valence and expectancy rather than the full valence model) were both positively associated with a student’s current study effort (Hypotheses 2a and 2b). These findings are in line with research that has found, depending on the circumstances (e.g., extensive experience with a decision), individuals can use variations of expectancy theory’s valence and force models that are simpler and thereby do not involve exhaustive information processing (e.g., [47]). These results are also consistent with Van Eerde and Thierry’s [13] meta-analysis, which recommended the use of specific expectancy

theory constructs (e.g., valence, expectancy) rather than the valence and force models, and with studies that have found empirical support for the validity of the theory’s individual constructs (e.g., [12, 99, 100]).

Based on our findings, neither the valence model nor the force model predicted a student’s final course grades as Hypotheses 3a and 3b were not supported. Given these results, neither model explained incremental variance in grades beyond cognitive ability as Hypotheses 4a and 4b were also not supported. Our findings are consistent with other studies in higher education that have examined expectancy theory’s validity for academic performance and observed nonsignificant effects and no incremental validity beyond cognitive ability (e.g., [41]). Indeed, these findings are in line with the overall pattern of results in these studies that have found weak support for the theory’s validity for academic success. Our study, however, extends previous investigations by conducting a more rigorous and appropriate assessment of expectancy theory’s validity based on Vroom’s [1] conceptualizations of the theory’s constructs and related methodological enhancements.

Our results provide further support for the most elementary version of expectancy theory as valence and the simplified force model were positively associated with a student’s final course grades (Hypotheses 5a and 5b). Furthermore, both valence and the simplified force model explained incremental variance in grades beyond cognitive ability as Hypotheses 6a and 6b were fully supported. These findings are also consistent with and further support the proposition that, at times, individuals will employ simpler versions of the theory’s models that

do not require extensive information processing [47] and thereby provide greater justification for the use of specific expectancy theory constructs rather than component models as recommended by Van Eerde and Thierry [13]. Additionally, these results are similar to findings in studies that have found support for the validity of the theory's individual constructs with performance as the criterion (e.g., [101]). Our study findings, however, extend past studies in this domain by being the first study, to the best of our knowledge, that has measured expectancy theory's constructs based on Vroom's [1] original conceptualizations and found evidence that motivation, although measured by a simplified force model, has incremental validity over cognitive ability in predicting academic performance. It should be noted that the simplified force model is conceptually most similar to the expectancy-value model of achievement motivation [25–27], which has been more consistently supported in higher education with academic performance as the criterion (e.g., [32, 33]).

Theoretical implications

Based on the results of studies in higher education that have explored the predictive validity of expectancy theory for performance (e.g., [38]) and its incremental validity beyond cognitive ability (e.g., [41]), there has been weak support for the theory given the observed low levels of validity for performance-related outcomes. Thus, it is reasonable to question the theory's applicability and generalizability to the college student population. However, our study, aimed to demonstrate that the weak support for expectancy theory in past studies can be partly explained by improper measurement of the theory's variables and other methodological limitations that were described in the Introduction section. As such, our study was intended to enable an appropriate assessment of the theory's validity.

Based on our findings, expectancy theory's valence and force models do not predict academic performance and thereby do not explain incremental variance beyond cognitive ability. However, the simplified force does predict grades and has incremental validity over cognitive ability. Therefore, although we are unable to find support for the validity of the theory's traditional force model, our results support the simplified force model (i.e., motivation = expectancy × valence) and thereby lend more credence to expectancy theory's validity for performance in higher education. Further, this finding provides added support to the argument that despite Vroom's [1] original formulation of the theory (i.e., intended to predict choice, intention, and effort), a framework that purports to assess motivation should also predict performance in achievement settings given the voluminous evidence that

has established motivation as one of two primary determinants of performance (e.g., [51, 66, 82]).

Arguably, the more important finding from our study is that the simplified force model predicts incremental variance in grades beyond cognitive ability. This is significant because it demonstrates that, contrary to past work, motivation, as assessed by an expectancy theory framework, can indeed explain unique variance in academic performance beyond cognitive factors when the common documented methodological limitations have been addressed. This finding is also significant because it highlights the predictive potential of a simplified force model over the traditional force model. Because our study enables a comparison of two different force models, one interpretation of this result is to view it as evidence of the greater validity of the simpler model over the traditional model. This finding can also be viewed as demonstrating the greater generalizability of the simplified force model as the version involving less comprehensive cognitive processing is found to apply to this sample as well as demonstrating the likelihood that the extensive processing requirements of the traditional force model may restrict its applicability to a limited set of conditions.

Based on the analysis of expectancy theory's predictive validity, another significant implication is regarding the stability of the simplified force model's validity coefficients across time (see Fig. 1). Apart from two weeks that can be considered outliers (Week 5 and 8), the validity coefficients are relatively stable and there is not a discernible difference between the strength of the relationship between motivation and grades when comparing the initial few weeks and the last few weeks. This trend reflects the predictive power of the simplified force model over the course of our study and is counterintuitive as it would be expected that the measured level of motivation that is more proximal to an outcome would have a stronger relationship with an outcome as compared to the measured level of motivation that is more distal to an outcome.

Lastly, based on our results, there was no observed difference in the validity of expectancy theory models with a multiplicative versus an additive integration of variables. When significant results were observed, they were found for both additive and multiplicative models. Alternatively, when results were not significant, they were observed for both additive and multiplicative models. Therefore, although an additive processing model had been observed to be more accurate for most participants [3], we found that with regard to statistical significance, the models were similarly valid and interchangeable. Readers are cautioned, however, against generalizing these findings as the type of information processing that individuals employ has been found to reflect individual

differences [2] and thereby these results might simply reflect characteristics of our sample.

Practical implications

Our results also have implications for counselors, academic advisors, and educators in higher education. The most direct practical implication would be for career counselors or academic advisors in the context of conducting academic advisement. Based on the significant finding for the simplified force model, an academic advisor can quickly attempt to gauge a student's study motivation for either currently enrolled courses or for the following term's courses that they plan to complete. This assessment will not require a lengthy survey. As described in our study, expectancy and valence (the direct measure) can both be assessed with just one item, which can then allow an advisor to rapidly determine the level of motivation for each course (i.e., study motivation for any one course = expectancy \times valence). Additionally, this assessment will enable an advisor to immediately diagnose the cause of the relatively low motivation for a course (assuming it applies to at least one course) by identifying whether (1) expectancy, (2) valence, or (3) both expectancy and valence are low.

Another practical implication would be in the context of providing undergraduate career counseling. Upon identifying the reason for low study motivation in a course, a counselor can work with a student to determine the proper intervention strategy that can enhance motivation and thereby increase academic performance. Depending on whether expectancy or valence is the concern, a counselor can then recommend the optimal solution(s) that a student can adopt. For example, if the deficiency in study motivation is due to low expectancy, one possible option would be for the student to seek supplemental instruction (e.g., tutoring) as task-specific knowledge and skills are proximal antecedents of expectancy [102]. A student who believes they possess the requisite level of course-specific skills and abilities will have a higher expectancy for a course [103, 104].

Alternatively, if the deficiency in study motivation is due to low valence, a potential recommendation would involve focusing on the likely rewards that are associated with excelling in a course and thus discussing the various intrinsic benefits (e.g., increased pride, higher competency in the subject) and extrinsic benefits (e.g., higher GPA, greater chance of qualifying for an internship that uses GPA as criteria) that the student can experience if performing at a high level [102]. Another viable option would be for the counselor to assist the student with drawing connections between what they are learning (or likely to learn soon) and their life [103]. When students

can clearly understand the relevance of a class to their lives, the valence of that class will be enhanced [105, 106].

Limitations and directions for future research

Our study has several limitations. First, because all study variables were based on self-reports by participants (i.e., single-source data), there is a concern with common method bias as observed relationships between variables might be inflated [107]. It is important to highlight, however, that student motivation is primarily assessed by self-reports as they are generally considered the most valid option [108]. Regardless, given the general concerns with common method bias when a study is based on single-source data, future research could take advantage of the various alternative approaches to assessing motivation (e.g., neuropsychological, physiological) by combining subjective self-report measures of motivation with these more objective approaches to triangulate the data. Related to the first limitation, a second limitation was that SAT scores and academic performance were also based on self-reports. Although self-reported grades and standardized test scores are valid and reliable substitutes when accessing scores from school records is not feasible [77], they can also be biased [84]. As noted by Kuncel et al. [77], self-reported grades are not likely to accurately reflect the performance of students who have low grades and thereby should be used with caution. We would, therefore, recommend future studies to seek institutional data, if possible, rather than rely on self-reported standardized test scores and grades. Another limitation of our study concerns external validity. Since our sample included college students from the United States, we would caution against generalizing study findings to students in different countries. Although past research has examined the generalizability of expectancy theory across cultures (e.g., [7, 16]), it has generally not included academic performance as an outcome and focused exclusively on the traditional force model. Therefore, future research is needed to explore the predictive validity of variations of the force model for academic performance across cultures. A fourth limitation of our study was the demographics of the sample, which included primarily first-year students (over 60%) and participants who identified as female (over 62%). Thus, caution is advised as our findings might not generalize to other populations. Besides motivation and ability, study skills and study habits are particularly important to academic performance [51]. Because the transition from high school to college involves adjusting to several academic demands and challenges [109], first-year students, who are in the earlier stages of this transition, are less likely to have learned the necessary study skills and developed the requisite study habits as compared with more experienced students.

Thus, study motivation might presumably be a better predictor of academic success for more experienced students who have already acquired the needed study skills and habits. We would, therefore, recommend future research to assess the validity of expectancy theory across different grade levels.

Conclusion

Because of various methodological concerns and inconsistent findings in studies that have examined the validity of expectancy theory, our study was motivated to assess the theory's validity for academic performance in a manner that was congruent with Vroom's [1] framework. Arguably, the most important finding in our study is that, unlike the traditional force model, the simplified force model predicts course grades and explains incremental variance beyond cognitive ability. Furthermore, contrary to what would be expected, the predictive validity of this model is relatively stable across the 11 weeks of the study. Given that our findings have several significant implications for expectancy theory as well as various practical implications, we believe they will motivate avenues of future research on the theory's validity in achievement settings.

Abbreviations

ALOC	Academic locus of control
ICC	Intraclass correlation coefficient
MLM	Multilevel modeling

Acknowledgements

Not applicable.

Authors' contributions

VP wrote the manuscript. VP and TS analyzed and interpreted the data. Both authors also read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This study was performed in accordance with the Declaration of Helsinki. Ethics approval was granted by the State University of New York at Albany's Institutional Review Board (#12208). Informed consent to participate in the study was required and obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 20 March 2024 Accepted: 31 July 2024
Published online: 13 August 2024

References

- Vroom VH. Work and motivation. Wiley; 1964.
- Donovan JJ. Work motivation. In: Anderson N, Ones DS, Sinangil HK, Viswesvaran C. Handbook of industrial, work and organizational psychology, Vol. 2. Organizational psychology. Sage;2002. pp. 53–76.
- Miner JB. Organizational behavior 1: essential theories of motivation and leadership. Sharpe ME. 2005.
- Tolman EC. Purposive behavior in animals and men. Century. 1932.
- Tolman EC. Principles of purposive behavior. In: Koch S. Psychology: a study of a science (Vol. 2). McGraw-Hill; 1959. pp. 92–157.
- Baker DD, Ravichandran R, Randall DM. Exploring contrasting formulations of expectancy theory. *Decis Sci*. 1989;20(1):1–13. <https://doi.org/10.1111/j.1540-5915.1989.tb01393.x>.
- Geiger MA, Cooper EA. Using expectancy theory to assess student motivation. *Issues Account Educ*. 1996;11(1):113–29.
- Harrell A, Stahl M. Additive information processing and the relationship between expectancy of success and motivational force. *Acad Manag J*. 1986;29(2):424–33. <https://doi.org/10.2307/256197>.
- Stahl MJ, Harrell AM. Modeling effort decisions with behavioral decision theory: toward an individual differences model of expectancy theory. *Organ Behav Hum Perform*. 1981;27(3):303–25. [https://doi.org/10.1016/0030-5073\(81\)90026-X](https://doi.org/10.1016/0030-5073(81)90026-X).
- Min H, Tan PX, Kamioka E, Sharif KY. Enhancement of study motivation model by introducing expectancy theory. *Int J Learn*. 2020;6(1):28–32. <https://doi.org/10.18178/IJLT.6.1.28-32>.
- Nebeker DM, Mitchell TR. Leader behavior: an expectancy theory approach. *Organ Behav Hum Perform*. 1974;11(3):355–67. [https://doi.org/10.1016/0030-5073\(74\)90025-7](https://doi.org/10.1016/0030-5073(74)90025-7).
- Pritchard RD, Sanders MS. The influence of valence, instrumentality, and expectancy on effort and performance. *J Appl Psychol*. 1973;57(1):55–60. <https://doi.org/10.1037/h0034197>.
- Van Eerde W, Thierry H. Vroom's expectancy models and work-related criteria: a meta-analysis. *J Appl Psychol*. 1996;81(5):575–86. <https://doi.org/10.1037/0021-9010.81.5.575>.
- Wanous JP, Keon TL, Latack JC. Expectancy theory and occupational/organizational choices: a review and test. *Organ Behav Hum Perform*. 1983;32(1):66–86. [https://doi.org/10.1016/0030-5073\(83\)90140-X](https://doi.org/10.1016/0030-5073(83)90140-X).
- Barba-Sánchez V, Atienza-Sahuquillo C. Entrepreneurial motivation and self-employment: evidence from expectancy theory. *Int Entrep Manag J*. 2017;13(4):1097–115. <https://doi.org/10.1007/s11365-017-0441-z>.
- Geiger MA, Cooper EA, Hussain I, O'Connell BT, Power J, Raghu-nandan K, Rama DV, Sanchez G. Using expectancy theory to assess student motivation: an international replication. *Issues Account Educ*. 1998;13(1):139–56.
- Lokman A, Hassan F, Ustadi YA, Rahman FAA, Zain ZM, Rahmat NH. Investigating motivation for learning via Vroom's theory. *Int J Acad Res Business Soc Sci*. 2022;12(1):504–30.
- Snead KC, Johnson WA, Ndede-Amadi AA. Expectancy theory as the basis for activity-based costing systems implementation by managers. In: Epstein MJ, Lee JY. *Advances in Management Accounting* (Vol. 14). Emerald Group Publishing Limited; 2005. pp. 253–275. [https://doi.org/10.1016/S1474-7871\(05\)14012-X](https://doi.org/10.1016/S1474-7871(05)14012-X).
- Tyagi PK. Diagnosing learning motivation of marketing students: An approach based on expectancy theory. *J Mark Educ*. 1985;7(2):28–34. <https://doi.org/10.1177/027347538500700205>.
- Ambrose ML, Kulik CT. Old friends, new faces: Motivation research in the 1990s. *J Manag*. 1999;25(3):231–92. <https://doi.org/10.1177/014920639902500302>.
- Campbell JP, Pritchard RD. Motivation theory in industrial and organizational psychology. In: Dunnette MD, editor. *Handbook of industrial and organizational psychology*. Rand-McNally; 1976. p. 63–130.
- Mitchell TR. Expectancy models of job satisfaction, occupational preference and effort: a theoretical, methodological, and empirical appraisal. *Psychol Bull*. 1974;81(12):1053–77. <https://doi.org/10.1037/h0037495>.

23. Pinder CC. Valence-instrumentality-expectancy theory. In: Steers RM, Porter LW. *Motivation and work behavior* (4th ed). McGraw Hill; 1987. pp. 69–89.
24. Schwab DP, Olian-Gottlieb JD, Heneman HG. Between-subjects expectancy theory research: a statistical review of studies predicting effort and performance. *Psychol Bull.* 1979;86(1):139–47. <https://doi.org/10.1037/0033-2909.86.1.139>.
25. Eccles JS, Adler TF, Futterman R, Goff SB, Kaczala CM, Meece JL, Midgley C. Expectancies, values, and academic behaviors. In: Spence JT, editor. *Achievement and achievement motivation: Psychological and sociological approaches*. Freeman; 1983. p. 75–146.
26. Eccles JS, Wigfield A. Motivational beliefs, values, and goals. *Annu Rev Psychol.* 2002;53(1):109–32. <https://doi.org/10.1146/annurev.psych.53.100901.135153>.
27. Wigfield A, Eccles JS. Expectancy–value theory of achievement motivation. *Contemp Educ Psychol.* 2000;25(1):68–81. <https://doi.org/10.1006/ceps.1999.1015>.
28. Atkinson JW. Motivational determinants of risk-taking behavior. *Psychol Rev.* 1957;64(6, Pt.1):359–72. <https://doi.org/10.1037/h0043445>.
29. Atkinson JW. An introduction to motivation. Van Nostrand; 1964.
30. Johnson ML, Taasobshirazi G, Clark L, Howell L, Breen M. Motivations of traditional and nontraditional college students: from self-determination and attributions, to expectancy and values. *J Contin High Educ.* 2016;64(1):3–15. <https://doi.org/10.1080/07377363.2016.1132880>.
31. Part R, Perera HN, Marchand GC, Bernacki ML. Revisiting the dimensionality of subjective task value: towards clarification of competing perspectives. *Contemp Educ Psychol.* 2020;62:101875. <https://doi.org/10.1016/j.cedpsych.2020.101875>.
32. Perez T, Dai T, Kaplan A, Cromley JG, Brooks WD, White AC, Mara KR, Balsai MJ. Interrelations among expectancies, task values, and perceived costs in undergraduate biology achievement. *Learn Individ Differ.* 2019;72:26–38. <https://doi.org/10.1016/j.lindif.2019.04.001>.
33. Rosenzweig EQ, Wigfield A, Hulleman CS. More useful or not so bad? Examining the effects of utility value and cost reduction interventions in college physics. *J Educ Psychol.* 2020;112(1):166–82. <https://doi.org/10.1037/edu0000370>.
34. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process.* 1991;50(2):179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T).
35. Ajzen I. *Attitudes, personality, and behavior* (2nd ed.). Open University Press; 2005.
36. Matsui T, Ohtsuka Y. Within-person expectancy theory predictions of supervisory consideration and structure behavior. *J Appl Psychol.* 1978;63(1):128–31. <https://doi.org/10.1037/0021-9010.63.1.128>.
37. Matsui T, Ikeda H. Effectiveness of self-generated outcomes for improving prediction in expectancy theory research. *Organ Behav Hum Perform.* 1976;17(2):289–98. [https://doi.org/10.1016/0030-5073\(76\)90068-4](https://doi.org/10.1016/0030-5073(76)90068-4).
38. Harrell A, Caldwell C, Doty E. Within-person expectancy theory predictions of accounting students' motivation to achieve academic success. *Account Rev.* 1985;60(4):724–35.
39. Youssef AA. Predicting student's effort and performance in foreign language courses: an application of expectancy theory of motivation [Paper presentation]. Detroit, MI: Teachers of English to Speakers of Other Languages 15th Annual Meeting; 1981.
40. Malloch DC, Michael WB. Predicting student grade point average at a community college from Scholastic Aptitude Tests and from measures representing three constructs in Vroom's expectancy theory model of motivation. *Educ Psychol Measur.* 1981;41(4):1127–35. <https://doi.org/10.1177/001316448104100422>.
41. Pringle CD. Expectancy theory: its applicability to student academic performance. *Coll Stud J.* 1995;29:249–55.
42. Katzell RA, Thompson DE. An integrative model of work attitudes, motivation, and performance. *Hum Perform.* 1990;3(2):63–85. https://doi.org/10.1207/s15327043hup0302_1.
43. Pinder CC. *Work motivation in organizational behavior* (2nd ed.). Psychology Press; 2008.
44. Vinacke WE. Motivation as a complex problem. In: Jones MR. *Nebraska symposium on motivation* (Vol. 10). University of Nebraska Press; 1962. pp. 1–45.
45. Polczynski JJ, Shirland LE. Expectancy theory and contract grading combined as an effective motivational force for college students. *J Educ Res.* 1977;70(5):238–41. <https://doi.org/10.1080/00220671.1977.10884996>.
46. Lawler EE. *Motivation in work organizations*. Brooks/Cole; 1973.
47. Lord RG, Hanges PJ, Godfrey EG. Integrating neural networks into decision-making and motivational theory: rethinking VIE theory. *Can Psychol.* 2003;44(1):21–38. <https://doi.org/10.1037/h0085815>.
48. Borman WC, White LA, Pulakos ED, Oppler SH. Models of supervisory job performance ratings. *J Appl Psychol.* 1991;76(6):863–72. <https://doi.org/10.1037/0021-9010.76.6.863>.
49. Campbell JP, McCloy RA, Oppler SH, Sager CE. A theory of performance. In: Schmitt N, Borman WC, editors. *Personnel selection in organizations*. Jossey-Bass; 1993. p. 35–70.
50. Chan D, Schmitt N, DeShon RP, Clause CS, Delbridge K. Reactions to cognitive ability tests: the relationships between race, test performance, face validity perceptions, and test-taking motivation. *J Appl Psychol.* 1997;82(2):300–10. <https://doi.org/10.1037/0021-9010.82.2.300>.
51. Credé M, Kuncel NR. Study habits, skills, and attitudes: the third pillar supporting collegiate academic performance. *Perspect Psychol Sci.* 2008;3(6):425–53. <https://doi.org/10.1111/j.1745-6924.2008.00089.x>.
52. Haertel GD, Walberg HJ, Weinstein T. Psychological models of educational performance: a theoretical synthesis of constructs. *Rev Educ Res.* 1983;53(1):75–91. <https://doi.org/10.2307/1170327>.
53. Jiang K, Lepak DP, Hu J, Baer JC. How does human resource management influence organizational outcomes? A meta-analytic investigation of mediating mechanisms. *Acad Manag J.* 2012;55(6):1264–94. <https://doi.org/10.5465/amj.2011.0088>.
54. Kanfer R, Chen G, Pritchard RD. *Work motivation: past, present, and future*. Routledge; 2008. <https://doi.org/10.4324/9780203809501>.
55. Katou AA, Budhwar PS. Causal relationship between HRM policies and organisational performance: evidence from the Greek manufacturing sector. *Eur Manag J.* 2010;28(1):25–39. <https://doi.org/10.1016/j.emj.2009.06.001>.
56. Maier NRF. *Psychology in industry: a psychological approach to industrial problems* (2nd ed.). Houghton Mifflin Company; 1955.
57. Mesmer-Magnus J, Viswesvaran C. Inducing maximal versus typical learning through the provision of a pretraining goal orientation. *Hum Perform.* 2007;20(3):205–22. <https://doi.org/10.1080/08959280701333016>.
58. O'Reilly CA III, Chatman JA. Working smarter and harder: a longitudinal study of managerial success. *Adm Sci Q.* 1994;39(4):603–27. <https://doi.org/10.2307/2393773>.
59. Diefendorff JM, Chandler MM. Motivating employees. In: Zedeck S. *APA handbook of industrial and organizational psychology, Vol. 3. Maintaining, expanding, and contracting the organization*. American Psychological Association; 2011. pp. 65–135 <https://doi.org/10.1037/12171-003>.
60. Kanfer R, Chen G. Motivation in organizational behavior: history, advances and prospects. *Organ Behav Hum Decis Process.* 2016;136:6–19. <https://doi.org/10.1016/j.jobhdp.2016.06.002>.
61. Latham GP, Pinder CC. Work motivation theory and research at the dawn of the twenty-first century. *Annu Rev Psychol.* 2005;56:485–516. <https://doi.org/10.1146/annurev.psych.55.090902.142105>.
62. Hattie J. *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge; 2009.
63. Kriegbaum K, Becker N, Spinath B. The relative importance of intelligence and motivation as predictors of school achievement: a meta-analysis. *Educ Res Rev.* 2018;25:120–48. <https://doi.org/10.1016/j.edurev.2018.10.001>.
64. Robbins SB, Lauver K, Le H, Davis D, Langley R, Carlstrom A. Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol Bull.* 2004;130(2):261–88. <https://doi.org/10.1037/0033-2909.130.2.261>.
65. Roth B, Becker N, Romeyke S, Schäfer S, Domnick F, Spinath FM. Intelligence and school grades: a meta-analysis. *Intelligence.* 2015;53:118–37. <https://doi.org/10.1016/j.intell.2015.09.002>.
66. Schneider M, Preckel F. Variables associated with achievement in higher education: a systematic review of meta-analyses. *Psychol Bull.* 2017;143(6):565–600. <https://doi.org/10.1037/bul0000098>.

67. Steinmayr R, Spinath B. The importance of motivation as a predictor of school achievement. *Learn Individ Differ*. 2009;19(1):80–90. <https://doi.org/10.1016/j.lindif.2008.05.004>.
68. Chamorro-Premuzic T, Harlaar N, Greven CU, Plomin R. More than just IQ: a longitudinal examination of self-perceived abilities as predictors of academic performance in a large sample of UK twins. *Intelligence*. 2010;38(4):385–92. <https://doi.org/10.1016/j.intell.2010.05.002>.
69. Freudenthaler HH, Spinath B, Neubauer AC. Predicting school achievement in boys and girls. *Eur J Pers*. 2008;22(3):231–45. <https://doi.org/10.1002/per.678>.
70. Kriegbaum K, Jansen M, Spinath B. Motivation: a predictor of PISA's mathematical competence beyond intelligence and prior test achievement. *Learn Individ Differ*. 2015;43:140–8. <https://doi.org/10.1016/j.lindif.2015.08.026>.
71. Lavrijsen J, Vansteenkiste M, Boncquet M, Verschuere K. Does motivation predict changes in academic achievement beyond intelligence and personality? A multitheoretical perspective. *J Educ Psychol*. 2022;114(4):772–90. <https://doi.org/10.1037/edu0000666>.
72. Spinath B, Spinath FM, Harlaar N, Plomin R. Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*. 2006;34(4):363–74. <https://doi.org/10.1016/j.intell.2005.11.004>.
73. Spinath B, Harald Freudenthaler H, Neubauer AC. Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Pers Individ Dif*. 2010;48(4):481–6. <https://doi.org/10.1016/j.paid.2009.11.028>.
74. Steinmayr R, Bipp T, Spinath B. Goal orientations predict academic performance beyond intelligence and personality. *Learn Individ Dif*. 2011;21(2):196–200. <https://doi.org/10.1016/j.lindif.2010.11.026>.
75. Weber HS, Lu L, Shi J, Spinath FM. The roles of cognitive and motivational predictors in explaining school achievement in elementary school. *Learn Individ Dif*. 2013;25:85–92. <https://doi.org/10.1016/j.lindif.2013.03.008>.
76. Trautwein U, Marsh HW, Nagengast B, Lüdtke O, Nagy G, Jonkmann K. Probing for the multiplicative term in modern expectancy-value theory: a latent interaction modeling study. *J Educ Psychol*. 2012;104(3):763–77. <https://doi.org/10.1037/a0027470>.
77. Kuncel NR, Credé M, Thomas LL. The validity of self-reported grade point averages, class ranks, and test scores: a meta-analysis and review of the literature. *Rev Educ Res*. 2005;75(1):63–82. <https://doi.org/10.3102/00346543075001063>.
78. Frey MC, Detterman DK. Scholastic assessment or g? the relationship between the scholastic assessment test and general cognitive ability. *Psychol Sci*. 2004;15(6):373–8. <https://doi.org/10.1111/j.0956-7976.2004.00687.x>.
79. Gottfredson LS, Crouse J. Validity versus utility of mental tests: example of the SAT. *J Vocat Behav*. 1986;29(3):363–78. [https://doi.org/10.1016/0001-8791\(86\)90014-X](https://doi.org/10.1016/0001-8791(86)90014-X).
80. Hunter JE. Cognitive ability, cognitive aptitude, job knowledge, and job performance. *J Vocat Behav*. 1986;29(3):340–62. [https://doi.org/10.1016/0001-8791\(86\)90013-8](https://doi.org/10.1016/0001-8791(86)90013-8).
81. Koenig KA, Frey MC, Detterman DK. ACT and general cognitive ability. *Intelligence*. 2008;36(2):153–60. <https://doi.org/10.1016/j.intell.2007.03.005>.
82. Richardson M, Abraham C, Bond R. Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol Bull*. 2012;138(2):353–87. <https://doi.org/10.1037/a0026838>.
83. Cole JS, Gonyea RM. Accuracy of self-reported SAT and ACT test scores: Implications for research. *Res High Educ*. 2010;51(4):305–19. <https://doi.org/10.1007/s11162-009-9160-9>.
84. Mayer RE, Stull AT, Campbell J, Almeroth K, Bimber B, Chun D, Knight A. Overestimation bias in self-reported SAT scores. *Educ Psychol Rev*. 2007;19(4):443–54. <https://doi.org/10.1007/s10648-006-9034-z>.
85. Batlis NC. Relationships between locus of control and instrumentality theory predictor of academic performance. *Psychol Rep*. 1978;43(1):239–45. <https://doi.org/10.2466/pr0.1978.43.1.239>.
86. Lied TR, Pritchard RD. Relationships between personality variables and components of the expectancy-valence model. *J Appl Psychol*. 1976;61(4):463–7. <https://doi.org/10.1037/0021-9010.61.4.463>.
87. Szilagyi AD, Sims HP. Locus of control and expectancies across multiple occupational levels. *J Appl Psychol*. 1975;60(5):638–40. <https://doi.org/10.1037/h0077156>.
88. Broedling LA. Relationship of internal-external control to work motivation and performance in an expectancy model. *J Appl Psychol*. 1975;60(1):65–70. <https://doi.org/10.1037/h0076353>.
89. Johnson RE, Rosen CC, Chang C-H, Lin S-H. Getting to the core of locus of control: Is it an evaluation of the self or the environment? *J Appl Psychol*. 2015;100(5):1568–78. <https://doi.org/10.1037/apl0000011>.
90. Galvin BM, Randel AE, Collins BJ, Johnson RE. Changing the focus of locus (of control): a targeted review of the locus of control literature and agenda for future research. *J Organ Behav*. 2018;39(7):820–33. <https://doi.org/10.1002/job.2275>.
91. Spector PE. Behavior in organizations as a function of employee's locus of control. *Psychol Bull*. 1982;91(3):482–97. <https://doi.org/10.1037/0033-2909.91.3.482>.
92. Trice AD. An academic locus of control scale for college students. *Percept Mot Skills*. 1985;61(3, Suppl):1043–6. <https://doi.org/10.2466/pms.1985.61.3f.1043>.
93. Raudenbush SW, Bryk AS. Hierarchical linear models (2nd ed). Sage; 2002.
94. Nezlek JB. Multilevel modeling for psychologists. In H. Cooper PM, Camic DL, Long AT, Panter D, Rindskopf, Sher KJ. *APA handbook of research methods in psychology, Vol. 3. Data analysis and research publication*. American Psychological Association; 2012. pp. 219–241. <https://doi.org/10.1037/13621-011>.
95. Bernerth JB, Aguinis H. A critical review and best-practice recommendations for control variable usage. *Pers Psychol*. 2016;69(1):229–83. <https://doi.org/10.1111/peps.12103>.
96. Hox J, Moerbeek M, Van de Schoot R. *Multilevel analysis: techniques and applications* (3rd ed.). Routledge; 2017.
97. Tabachnick BG, Fidell LS. *Using multivariate statistics* (7th ed.). Pearson; 2019.
98. Aguinis H, Gottfredson RK, Joo H. Best-practice recommendations for defining, identifying, and handling outliers. *Organ Res Methods*. 2013;16(2):270–301. <https://doi.org/10.1177/1094428112470848>.
99. Behling O, Dillard JF, Gifford WE. Tests of expectancy theory predictions of effort: a simulation study comparing simple and complex models. *J Bus Res*. 1979;7(4):331–47. [https://doi.org/10.1016/0148-2963\(79\)90011-0](https://doi.org/10.1016/0148-2963(79)90011-0).
100. Muchinsky PM. A comparison of within- and across-subjects analyses of the expectancy-valence model for predicting effort. *Acad Manag J*. 1977;20(1):154–8. <https://doi.org/10.2307/255470>.
101. Klein HJ. Further evidence on the relationship between goal setting and expectancy theories. *Organ Behav Hum Decis Process*. 1991;49(2):230–57. [https://doi.org/10.1016/0749-5978\(91\)90050-4](https://doi.org/10.1016/0749-5978(91)90050-4).
102. Klein HJ, Austin JT, Cooper JT. Goal choice and decision processes. In: Kanfer R, Chen G, Pritchard RD, editors. *Work motivation: Past, present, and future*. Routledge; 2008. p. 101–50.
103. Hulleman CS, Barron KE, Kosovich JJ, Lazowski RA. Student motivation: Current theories, constructs, and interventions within an expectancy-value framework. In Lipnevich AA, Preckel F, Roberts RD. *Psychosocial skills and school systems in the 21st century: Theory, research, and practice*. Springer International Publishing; 2016. pp. 241–278. https://doi.org/10.1007/978-3-319-28606-8_10.
104. Wigfield A, Eccles JS. *Development of achievement motivation*. Academic Press; 2002.
105. Hulleman CS, Godes O, Hendricks BL, Harackiewicz JM. Enhancing interest and performance with a utility value intervention. *J Educ Psychol*. 2010;102(4):880–95. <https://doi.org/10.1037/a0019506>.
106. Hulleman CS, Harackiewicz JM. Promoting interest and performance in high school science classes. *Science*. 2009;326(5958):1410–2. <https://doi.org/10.1126/science.1177067>.
107. Podsakoff PM, Podsakoff NP, Williams LJ, Huang C, Yang J. Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annu Rev Organ Psych Organ Behav*. 2024;11:17–61. <https://doi.org/10.1146/annurev-orgpsych-110721-040030>.

108. Fulmer SM, Frijters JC. A review of self-report and alternative approaches in the measurement of student motivation. *Educ Psychol Rev.* 2009;21(3):219–46. <https://doi.org/10.1007/s10648-009-9107-x>.
109. Permzadian V, Credé M. Do first-year seminars improve college grades and retention? A quantitative review of their overall effectiveness and an examination of moderators of effectiveness. *Rev Educ Res.* 2016;86(1):277–316. <https://doi.org/10.3102/0034654315584955>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.