

# Measuring cost: The forgotten component of expectancy-value theory



Jessica Kay Flake <sup>a,\*</sup>, Kenneth E. Barron <sup>b</sup>, Christopher Hulleman <sup>c</sup>, Betsy D. McCoach <sup>a</sup>, Megan E. Welsh <sup>a</sup>

<sup>a</sup> University of Connecticut, USA

<sup>b</sup> James Madison University, USA

<sup>c</sup> University of Virginia, USA

## ARTICLE INFO

### Article history:

Available online 9 March 2015

### Keywords:

Cost  
Expectancy-value model  
Scale development  
Motivation  
Measurement

## ABSTRACT

Although the Expectancy-Value Model offers one of the most influential models for understanding motivation, one component of this model, cost, has been largely ignored in empirical research. Fortunately, recent research is emerging on cost, but no clear consensus has emerged for operationalizing and measuring it. To address this shortcoming, we outline a comprehensive scale development process that builds and extends on prior work. We conducted a literature review of theory and existing measurement, a qualitative study with students, a content alignment with experts, exploratory and confirmatory factor analysis, and a correlational study. In the literature and across our studies, we found that cost was salient to students, separate from expectancy and value components, contained multiple dimensions, and related to student outcomes. This work led to proposing a new, 19 item cost scale with four dimensions: task effort cost, outside effort cost, loss of valued alternatives cost, and emotional cost. In addition, to extend existing cost measures, careful attention was taken to operationalize the cost dimensions such that the scale could be easily used with a wide variety of students in various contexts. Directions for future research and the implications for the study of motivation are discussed.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

When we asked students to describe characteristics of the class in which they were the least motivated, we heard the following responses: “It was just so much, I couldn’t cram everything into my head”, “Studied so much for this class that I had to sacrifice work for other classes”, “It was really stressful with all the work we had to do.”

What is it that these students are describing? Can it be measured systematically? How is it related to students’ motivation and academic performance? And, what could teachers do to optimize student motivation if they knew students were experiencing it?

Motivation science offers a number of different options for understanding student motivation (Pintrich, 2003). In the current paper, we turn to expectancy-value models (Eccles et al., 1983) to understand what the students quoted above have expressed. In particular, one component within the expectancy-value model, known as cost, captures what the students are describing. Eccles (2005) defined cost as “what an individual has to give up to do a task, as well as the anticipated effort one will need to put into task completion.” Although cost has been theorized as an important component of

the expectancy-value model, empirical work within the expectancy-value framework has historically neglected it (Wigfield & Cambria, 2010). Fortunately, a growing body of work is beginning to emerge (Barron & Hulleman, 2015; Chen & Liu, 2009; Chiang, Byrd, & Molin, 2011; Conley, 2012; Perez, Cromley, & Kaplan, 2014; Trautwein et al., 2012; Watkinson, Dwyer, & Nielsen, 2005). In the current paper, we review what is currently known about cost and explore how it is experienced by undergraduate students using qualitative methods. We then use this theoretical foundation to develop a new measure of cost and present initial validity evidence for the scale.

## 2. Review of the cost literature

Over thirty years ago, Eccles et al. (1983) were the first to translate expectancy-value models of motivation into educational research. This framework proposes that motivation is a function of expectancy (i.e., students’ perceived judgments of their ability to succeed) and task value (i.e., students’ perceived level of task importance) components. Cost was first introduced by Eccles et al. (1983) as a mediator that would impact an individual’s overall value for an activity. Specifically, cost was hypothesized to be influenced by three dimensions: perceived effort, loss of valued alternatives, and the psychological cost of failure. Perceived effort was described as students’ perception of how much effort is needed to be successful at the task, stating that cost will be high if that effort is not perceived to be worth the benefit. Loss of valued alternatives was hypothesized to occur

\* Corresponding author. Educational Psychology, Measurement, Evaluation, and Assessment, University of Connecticut. Neag School of Education, 249 Glenbrook Road, Unit 3064 Storrs, Connecticut 06269, USA.

E-mail address: [kayflake@gmail.com](mailto:kayflake@gmail.com) (J.K. Flake).

when engaging in one activity prevents an individual from being able to participate in other valued activities. Finally, the psychological cost of failure was described as the anxiety related to the potential of failure at the task. This initial conceptualization implies that cost is a negative motivational component that subtracts from the overall level of value a student has for the task.

The first attempts to measure cost are summarized in a 1980 grant report authored under Eccles's maiden name (see [Parsons et al., 1980](#)). In particular, in the subscale "Cost of Effort to Do Well in Math," items assessing whether the amount of effort is worthwhile and loss of valued alternatives were included (see [Table 1](#) for items), representing two of the three dimensions that [Eccles et al. \(1983\)](#) theorized as cost. However, subscales to measure task effort and task difficulty were also included that were theorized to impact a student's expectancy (see [Table 1](#) for items). Thus, in this initial work, it was difficult to disentangle the different appraisals of effort to distinguish one as cost (a component of value), one as difficulty (a component of expectancy), or one as general effort (separate from expectancy and value).

Eccles and colleagues used the [Parsons et al. \(1980\)](#) expectancy-value scale, or an adaptation of it, for years. Their published work focused on the predictive power of the expectancy and value subscales (but not the cost subscales), contributing greatly to what we know about student motivation. Then, in 1995, Eccles and Wigfield formally revisited the measurement properties of the scale by investigating the structure of expectancy and value, as well as the task difficulty items. The authors used exploratory factor analyses to refine the item pool and confirmatory factor analyses to investigate the structure of the different expectancy-value components. Although expectancy, value, and task difficulty were included in the analysis, the original cost subscale was not. Only one item from the cost subscale was featured in the analysis, but was included as an item on the attainment value subscale.

The results of [Eccles and Wigfield \(1995\)](#) provided evidence that: (a) expectancy and value components are distinct, though positively related, and (b) task difficulty is both separate from and differentially related to expectancy and value. Because the effort component of task difficulty is similar in content to the [Parsons et al. \(1980\)](#) "Cost of Effort to do Well in Math" items, cost might also be distinct from expectancy and value and differentially relate to them. Although Eccles and Wigfield offered an important and rigorous test of the factor structure of their scale and the interrelations between components, the relationship between cost and other constructs was not formally evaluated.

In 2000, Wigfield and Eccles offered additional clarification about the constructs of their expectancy-value model in a special issue of *Contemporary Educational Psychology* focused on clarifying motivation constructs. For cost, they continued to define the effort and loss of valued alternatives dimensions similarly to [Eccles et al. \(1983\)](#);

but rather than focusing on just psychological cost of failure or success, they offered a broader definition of "emotional cost" suggesting emotional costs could encompass other mood states. Further, rather than referring to cost as a mediator of value, as was done in [Eccles et al. \(1983\)](#), they present cost as a type of value, as was done in most writing on cost after 1983.

Over the past ten years, other educational psychologists have been inspired by Eccles's model to explore cost both qualitatively and quantitatively and its effects on student outcomes ([Battelle & Wigfield, 2003](#); [Chen & Liu, 2009](#); [Chiang et al., 2011](#); [Conley, 2012](#); [Perez et al., 2014](#); [Trautwein et al., 2012](#); [Watkinson et al., 2005](#)). The surge of cost-related work emphasizes the interest and need in the field for theoretical and measurement clarification. This emerging research also suggests that cost does contribute to student motivation and is separate from other value components. For example, [Watkinson et al. \(2005\)](#) conducted a qualitative study of elementary students' motivation to be physically active during recess. Students discussed the cost of engaging in activities without being prompted, including both the physical costs of engaging in an activity (fatigue or being cold) and also psychological or emotional costs (being teased by friends or facing scrutiny).

[Chen and Liu \(2009\)](#) also qualitatively studied cost, noting the lack of a psychometrically sound measure and deep understanding of cost. For example, they asked students the open ended question, "If you have a choice whether to take physical education, would you rather not take it or [would] you still want to take it, and why?" Those who responded that they would not take physical education cited other demands on their time and heavy workload as contributing factors. Interestingly, this highlights a new dimension of cost. While the amount of effort required by a given task has always been recognized as a dimension of cost, educational psychologists have not considered how effort needed for other tasks can increase feelings of cost. For example, students may experience higher levels of cost in a particular class because of the time and energy they need to spend on another class, or other activities. Having other competing demands and how it impacts human behavior is more widely discussed in other literatures such as behavioral economics (for review, see [Madden, 2000](#)).

Building from [Watkinson et al.'s \(2005\)](#) qualitative work, [Chiang et al. \(2011\)](#) studied the cost of engaging in physical activity quantitatively. They included one item to assess each of the cost dimensions proposed by Eccles, Wigfield, and their colleagues (see [Table 2](#) for items), along with items to measure expectancy and value. Exploratory factor analysis revealed that a two-factor structure best fits the data, with one factor including the three cost items, and the other including the items for expectancy, interest value, and importance value (which they labeled as "beliefs"). Students who reported higher levels of cost reported less exercise, whereas students who reported higher levels on the beliefs subscale reported

**Table 1**  
Cost and effort scale items from [Parsons et al. \(1980\)](#).

Cost of effort to do well in math	Effort	Difficulty of current math
1. Is the amount of effort it will take to do well in your math course this year worthwhile to you?	1. How hard to do you have to try get good grades in math?	1. In general, how hard is math for you?
2. Is the amount of effort it would take to do well in advanced high school math courses worthwhile to you?	2. How hard do you have to study for math tests to get a good grade?	2. Compared to most other students in your class, how hard is math for you?
3. How much does the amount of time you spend on math keep you from doing other things you would like to do?	3. To do well in math, I have to work: much harder in math than in other subjects to much harder in other subjects than in math	3. Compared to most other school subjects that you have taken or are taking, how hard is math for you?
	4. How much time do you spend on home work?: an hour or more to I rarely do any math homework	
	5. How hard do you try in math?	
	6. Compared to most other students you know, how much time do you have to spend working on your math assignments?	

*Note:* Response options for items were on a 7-point scale with anchors at the low and high extreme (e.g., 1 (not at all) to 7 (very much) for Cost of Effort to do Well in Math, 1 (very easy) to 7 (very hard) for Effort, and 1(very worthwhile) to 7 (not at all worthwhile) for Difficulty of Current Math).

**Table 2**

Partial cost scales from other educational psychologists.

**Chiang et al. (2011)**

1. When you exercise, how much are you missing out on doing other things?
2. How hard is it for you to exercise?
3. Does exercising make you feel worn out and tired, so you don't want to exercise again?

**Trautwein et al. (2012)**

1. I'd have to sacrifice a lot of free time to be good at mathematics/English.
2. I'd have to invest a lot of time to get good grades in mathematics/English.

**Conley (2012)**

1. I have to give up a lot to do well in math.
2. Success in math requires that I give up other activities I enjoy.

**Luttrell et al. (2010)**

1. Math exams scare me.
2. Trying to do math causes me a lot of anxiety.
3. Taking math classes scares me.
4. I worry about getting low grades in my math courses.
5. I have to study much harder for math than for other courses.
6. Mathematical symbols confuse me.
7. Solving math problems is too difficult for me.

more exercise. Additionally, a weak, negative relationship between the beliefs factor and cost also was found. However, scores on the cost subscale had low reliability ( $\alpha = .57$ ).

Conley (2012) and Trautwein et al. (2012) included brief, two item measures of cost in larger quantitative studies of motivation. Conley (2012) tested two loss of valued alternatives items, whereas Trautwein et al. (2012) tested one item to assess the amount of effort required and one item to evaluate the loss of valued alternatives (see Table 2 for items). In both studies, factor analyses revealed that cost items separated into a factor, which was distinct from the other hypothesized value scales and negatively related to expectancy and value scales. In both studies, cost was critical in discriminating which students had a more or less adaptive profile of motivation.

Finally, Luttrell et al. (2010) developed the Mathematics Value Inventory to measure each type of value proposed in Eccles's model: interest value, utility value, attainment value, and cost. Their cost scale included seven items. Four items (Table 2) assessed emotional consequences (such as anxiety), and three items assessed the difficulty of math. They found that students who took more math courses reported significantly lower levels of cost and higher levels of interest and utility value.

In contrast to brief scales, two efforts have been made to develop more comprehensive measures of cost (Battle & Wigfield, 2003; Luttrell et al., 2010; Perez et al., 2014). Battle and Wigfield (2003) tested 24 items to measure the anticipated cost or personal sacrifice associated with females' pursuit of a graduate degree. Five items assessed required effort in relation to the worth of graduate school, six items measured the loss of valued alternatives, six items described the psychological cost of failure, and the last seven items expressed ambivalence about the worth of pursuing graduate school based on the consequences. Exploratory factor analyses were "unsuccessful in identifying clear, distinct factors" (Battle & Wigfield, 2003, p. 63), with only 9 of the 24 cost items loading on a single factor. The remaining items were discarded due to cross loadings with other value subscales, resulting in a single factor, 9-item cost subscale. The cost subscale was negatively related to intentions to attend graduate school.

Perez et al. (2014) extended Battle and Wigfield's (2003) work by adapting 20 items to capture the cost of effort, loss of valued alternatives (which they labeled opportunity cost), and the psychological cost of majoring in a science-related field (see Table 3 for items). Unlike the factor solution found in Battle and Wigfield, exploratory factor analyses supported a three-factor solution for cost that coincided with the effort, loss of valued alternatives, and psychological subscales. The subscales were then used in path analyses and found to significantly predict intentions to leave a STEM major,

**Table 3**

Cost items from Perez, Cromley, and Kaplan (2014).

**Effort cost**

1. Considering what I want to do with my life, having a science major is just not worth the effort.
2. My science major would not be worth it if I had to work hard after graduating to repay a long-term tuition loan.
3. When I think about the hard work needed to get through my science major, I am not sure that getting a science degree is going to be worth it in the end.
4. Getting a science degree sounds like it really requires more effort than I'm willing to put in.
5. I worry that I will waste a lot of time and money before I find out that I do not want to continue my science major.\*
6. I am not sure if I've got the energy to work and be a science major at the same time.\*

**Opportunity cost (family and friends)**

7. I worry about losing track of some valuable friendships if I'm in a science major and my friends are not.
8. I'm concerned my science major may cost my some treasured friendships.
9. I'm concerned my science major will take time away from other activities that I want to pursue.
10. I'm concerned my science major may cause family relationships to suffer.
11. I worry that my science major will take time away from other activities that I want to pursue.
12. I'm concerned that my career goals in science will prevent me from being able to focus on marriage and family soon as I'd like to.
13. I'd rather leave more time for fun than for something as intense as a science major.\*
14. I do not want to take time away from a job and earning money by remaining in my science major.\*

**Psychological cost**

15. I'm concerned that I'm not a good enough student to do well in my science major.\*
16. My self-esteem would suffer if I tried my science major and was unsuccessful at it.
17. I would be embarrassed if I found out that my work in my science major was inferior to that of my peers.
18. I'm concerned that I won't be able to handle the stress that goes along with my science major.
19. It frightens me that the courses required for my science major are harder than courses required for other majors.
20. I'm concerned my science major may cause me to be viewed by other people as a person with peculiar goals or interests.\*

\* Items had low loadings or cross-loadings.

with the effort subscale having the strongest effect. Thus, Perez et al. (2014) offered the first evidence that the theorized dimensions of cost contribute to understanding student behavior differentially. A limitation for widespread adoption of this scale is that the cost items center on particular issues for college students (e.g., student loans) and a particular type of loss of valued alternatives (e.g., family and friends).

In sum, our literature review of work in educational psychology highlights that the construct of cost is salient to students, separate from other components of the expectancy-value model, and related to a number of educational outcomes. However, different approaches to defining and measuring cost underscore a need for further clarification. This can be seen when looking at the items present in Tables 1–3, as each study measures a different type of cost with different theoretical implications. In addition, a number of scales focus on cost issues for a particular activity (e.g., physical exercise) or a particular student population (e.g., college). Thus, additional scale work to develop multidimensional cost scales for use in wider contexts and with other student populations is needed.

### 3. Scale development framework

A key step in developing a measure is establishing its validity, which reflects the extent to which particular interpretations and uses are appropriate for that measure (American Educational Research Association, American Psychological Association, & National

Council on Measurement in Education, 2014). Because measures can be used in multiple ways, and because validation results can be used to continuously improve measures, validation is sometimes described as a never-ending process (Messick, 1980). We present preliminary validation evidence here. To do so, we use three validation phases described by Benson (1998): substantive, structural, and external, as they provide an accessible framework and guide to scale development. The purpose of the *substantive phase* is to evaluate how a construct is defined, operationalized, and measured. During the substantive phase, researchers examine theory and previous research to define the construct and also use empirical work to capture the full range of potential ways it may be operationalized. This phase aims to answer the first question posed earlier about the students' comments, "What is it that these students were describing?"

The purpose of the *structural phase* is to conduct a psychometric investigation of the scale that emerges from the substantive phase. Research conducted during the structural phase has an internal focus, where the scale's psychometric properties and internal consistency are scrutinized. This phase aims to answer the second question posed earlier, "Can it be measured systematically?" Finally, the *external phase* examines how the construct under investigation relates to other constructs in order to establish convergent, discriminant, and criterion-related validity evidence. This phase aims to answer the third question posed earlier, "How is it related to students' motivation and academic performance?"

The three construct validation phases are meant to proceed developmentally, starting with the substantive phase, moving to the structural phase, and concluding with the external phase. As a result, we have organized our paper into three separate sections to summarize the validity evidence collected for each phase.

#### 4. Substantive phase

To execute the substantive phase, we adopted a multi-step approach recommended by Gehlbach and Brinkworth (2011). First, we developed an initial conceptualization for the substance of cost (see Fig. 1) based on our literature review. This conceptualization is meant to be temporary, providing a framework for further exploring the substance of cost. We started with an overall, working definition for cost: *what is invested, required, or given up to engage in a task*, and we outlined four possible cost dimensions. Three dimensions were consistently discussed in the literature from Eccles's model: effort, loss of valued alternatives, and psychological/emotional. We

also found evidence for an additional dimension, *outside effort cost*, which was initially revealed in a qualitative investigation of cost by Chen and Liu (2009). This motivated us to consider literature outside of educational psychology, specifically behavioral economics, which describes behavior as a function of how much energy or time is required for all tasks in concert with one another (Madden, 2000). Therefore, we suggest the structure of cost include the effort exerted for other tasks to better understand what drives or inhibits student behavior for a given task. We defined outside effort cost as the time, energy or effort put forth for tasks other than the one of interest.

With this initial conceptualization as a guide, we conducted qualitative interviews with undergraduate students, and synthesized our findings with the literature review to converge on a final theoretical structure and operational definitions for cost. To complete the substantive phase, we generated an initial pool of items to measure the theoretical structure of cost, and revised them based on feedback from a panel of experts. Each of these steps is discussed in detail below.

##### 4.1. Focus groups with students

Per Gehlbach and Brinkworth's (2011) recommendation, we set out to further understand the construct of cost by conducting a qualitative investigation. We conducted focus groups to see if students experienced cost and discussed the four dimensions of cost uncovered from the literature review. We also examined whether students used similar terminology to what is present in the literature. Gehlbach and Brinkworth (2011) noted that researchers rarely use qualitative methods to examine constructs prior to item writing. The advantage of including these types of studies is that researchers can ascertain if their theoretical conceptualization matches how the population of interest thinks about it, an approach consistent with Benson's (1998) description of the substantive phase of validation, which includes both theoretical and empirical components. This approach is beneficial because it also allows the theoretical structure to be clarified and amended based on empirical results.

##### 4.1.1. Participants and procedure

One-hundred and twenty-three students from a medium sized, public university in the southeastern United States chose to participate in focus group interviews about their motivation for college classes. We used the university's participant pool to list our study and students chose to participate through the online system. Students are required to participate in research each semester, and they

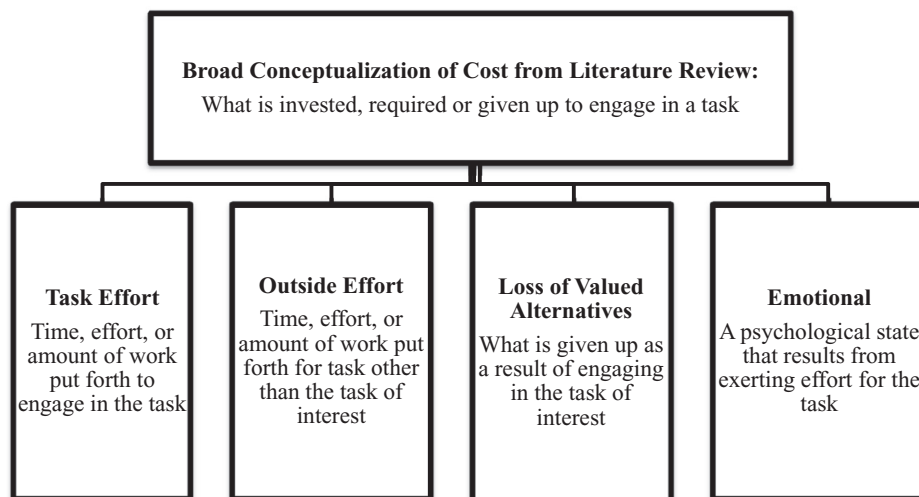


Fig. 1. Definition and specific theoretical dimensions of cost from literature review.



received research credit for participation in our study. Focus groups were conducted with a minimum of two people and a maximum of six people, and 29 separate focus groups were held. The sample was mostly female (71%) and Caucasian (89%) with an average age of 19.45 ( $SD = 2.80$ ).

Focus groups took place in a small conference room with two researchers present. One provided instructions and facilitated discussion; the other typed participants' responses into an electronic document that was projected on a screen for all to see. Each session followed a detailed protocol and script. We choose to conduct focus groups specifically because we wanted to take advantage of synergy and idea building among participants. Students were prompted to think of two types of experiences: a class in which they were the most motivated (MM) and then the class in which they were the least motivated (LM). First, we had them respond to our prompts individually in written form. Then they shared with the group, which allowed participants to respond to each other. The focus group approach gave students the opportunity to reflect individually at first and then to react and expand upon what others had said (Hesse-Biber & Leavy, 2011). Once a student shared his or her idea, the assisting researcher would type their idea verbatim, which was projected on screen for everyone to see. If students' responses were vague, the lead researcher asked a probing question. For example, "Can you elaborate a bit more on your idea so I know why it was motivating?" Care was taken not to lead respondents, but to only ask for more elaboration. The assisting researcher confirmed that each response was correctly typed before moving to the next student, who then shared a new idea. Students continued sharing ideas until no new ideas were forthcoming.

4.1.2. Overview of MM and LM coding

Coding of the unique, typed responses took a top-down approach. We used our working conceptualization of cost that came from the literature review (see Fig. 1) to create a coding rubric. The first two authors independently coded the responses using the rubric. The rubric outlined the four major categories for cost from Fig. 1: effort related to the task (i.e., class), effort unrelated to the task (i.e., outside cost), loss of valued alternatives, and emotional consequences. At this stage, we wanted to explore how students conceptualized their effort, so any description of effort was coded as a cost. This allowed us to further consider the valence and content of those descriptions. Further, though an initial coding scheme was used, it was possible that coders would find responses that were related to cost, but were not cleanly captured by the rubric's four categories. Appreciating the possibility that other themes could emerge from the data (Creswell, Hanson, Clark Plano, & Morales, 2007), the coders also kept track of any related responses that did not fit the rubric definitions.

Coders progressed through two levels of coding where they first considered if the response was a cost response, then which theoretical dimension of cost it described. For example, the response "The course was too intense, too much time, too rigorous", was first coded as *Cost*. Then during the second level of coding, it was coded as *Effort*. In contrast, the response, "I hate math", was coded as *Not Cost*. If a student mentioned two cost dimensions in one response, it was coded for both and counted as two instances of cost. If a student response was coded *Cost* but fell outside of the theorized dimensions, the coders coded it as *Other*. We then explored the *Other* category for emerging themes related to the construct of cost. Percentage agreement was calculated for the two coders, and any disagreements were resolved through additional discussion.

4.1.3. MM and LM results

After the first round of coding, percentage agreement for the first level of coding (*Cost* or *Not Cost*) was 83%. Specific disagreements were caused by one coder rating feelings of pressure to do well for

a class and descriptions of the class being too easy or not having enough required work as *Cost*, whereas the other did not. To clarify these discrepancies, the two coders met and discussed each point of disagreement. After this discussion, they decided that responses describing the pressure to do well were *Cost* responses, and those responses were recoded as *Cost* for the first level and *Emotional* for the second level. Also, the descriptions of the class being "too easy" or "not enough work" were considered unique and informative, so the coders created a new code, *Not Enough Cost* to categorize those responses. After these recodes, agreement for the first level of coding was 100%. At the second level of coding the coders were in 98% agreement. Any discrepancies for the second level of coding were resolved through discussion.

A total of 708 responses were coded across both the MM (321 responses) and LM parts (387 responses) of the qualitative study. Responses were not tied to a specific individual or focus group in this analysis. Instead, frequencies and percentages were calculated using all the unique, typed responses that were shared within each focus group. The percentages of Level 1 coding of *Not Cost* and *Cost*, and Level 2 coding for the specific components of cost are included for in Table 4.

The first noteworthy finding was that general descriptions of cost (when coded as what is exerted or given up to engage in a task) were present across both the MM and LM parts of the study. Overall, 15% of responses were coded as *Cost* in the MM class, and 11% of responses were coded as *Cost* in the LM class. Then, more specifically, responses initially coded as *Cost* were further coded for the theoretical dimensions of *Task Effort*, *Outside Effort*, *Loss of Valued Alternatives*, *Emotional*, or *Other*. An unexpected theme, *Not Enough Cost*, also emerged from the LM condition and accounted for an additional 3% of the responses. In these responses, students' indicated that they were not motivated in the class because it did not require them to put in a substantial amount of effort, time, or was "too easy." Though rare, these responses provided a unique insight into effort and should be considered in future research.

*Effort* was the most common cost dimension mentioned in both the MM and LM parts (representing 58% and 42% of the responses, respectively). However, when examining the content of these codes, we noticed vast differences. In the MM class, students described effort as "keeping up" or "quizzes forced me to study a lot", but in the LM class, effort was described as "too much." Thus, while the MM responses were positive in nature and cited as reasons why the students were motivated in that class, the LM responses were always negative.

After effort, the most frequent dimension in the MM class was *Other* (25%). The *Other* responses consisted of descriptions of the amount of credit hours the course was and needing to work hard because the course was a requirement for the major. Though these responses described effort, they did not map on to any of the cost dimensions. The final dimension found in the MM data was *Emotional* (17%). These responses consisted of descriptions of challenge and social pressure to do well (e.g., from family, friends, or the professor).

**Table 4**  
Percentage of responses coded as cost and cost subcomponents by class type.

	Most motivated class	Least motivated class
Level 1 Coding – Cost responses	15%	11%
Level 1 Coding – Not enough cost emerging theme	0%	3%
Level 2 Coding –Task–effort cost	58%	42%
Level 2 Coding – Loss of valued alternatives	0%	44%
Level 2 Coding – Emotional costs	17%	9%
Level 2 Coding – Other	25%	5%

After effort, the most frequent dimension in the LM class was *Loss of Valued Alternatives* (44% of *Cost* responses). These responses described giving up time, particularly time to sleep when classes were early in the morning. The remaining dimensions found in the LM class were *Emotional* (9%) and *Other* (5%). The *Emotional* responses described stress or negative feelings related to the class. The *Other* responses described financial cost, as well as other ideas that were not captured in the cost dimensions.

#### 4.1.4. Discussion of substantive phase

As Benson (1998) explains, the substantive phase includes both theoretical and empirical components, which must be synthesized to define the construct prior to writing items. Our literature review of cost by educational psychologists consistently found cost described as having three major dimensions: task effort cost, loss of valued alternatives, and emotional cost. In addition to the historical components, we proposed a new component, outside effort cost. This component provides an opportunity to better understand what drives or inhibits student behavior for a given task, after taking into consideration other factors in the student's life.

The existing research provided a theoretical foundation, but we gain more insight into the construct through our qualitative study. We learned that students discuss heavy workload and required effort when they are motivated and unmotivated. What varied across their most motivated and least motivated classes were the appraisals of that effort and how it made the students feel. Students described challenge and pressure to do well when they were most motivated. Additionally, effort in their most motivated class was described as a reason for their motivation. In contrast, in their least motivated class, students described their work as “too much”, “overwhelming”, or “stressful.”

These differences in the most motivated and least motivated responses emphasize the importance of the subjective appraisal of the effort invested and revealed how effort should be operationalized if we want to capture a cost and barrier to motivation. When effort is appraised negatively, cost is activated, and we term this dimension as task effort cost. Simply measuring the overall amount of effort or time a task requires alone cannot be used to distinguish the motivated student from the unmotivated student. A subjective appraisal of the work being *too much* must be captured. For example, a cost item that reads, “This class requires a lot of effort” may be endorsed by students who are motivated by the challenge and by

students who are overwhelmed. In contrast, a cost item that reads, “This class requires too much effort” more clearly captures an appraisal that would negatively impact motivation.

Similarly, we learned that the loss of valued alternatives is only salient to college students in their least motivating classes. Students did not describe giving up or sacrificing other valued activities in their most motivating classes. Thus, a negative appraisal (like experiencing feelings of sacrifice) again appears to be critical to activate cost. For these reasons, we operationalize cost as negative appraisals of effort, negative appraisals of effort required from other tasks, loss of valued alternatives, and negative appraisals of one's emotional state. With the operational definitions included in Fig. 2, we wrote 45 candidate items, more than we would want on the final scale. We recognized that some items might be unclear, function poorly, or not capture the construct in the way we intended.

## 4.2. Content alignment

To complete the substantive phase, we presented our initial pool of items to a panel of experts in motivation theory to collect content validity evidence (American Educational Research Association et al., 2014). Eight content experts participated in a content alignment (i.e., backward translation, Anderson & Thelk, 2005; Dawis, 1987; Smith & Kendall, 1963). This method calls for the use of a group of trained judges or experts to map items to the theorized scale structure. Experts also offered feedback about the wording and relevance of each item. We used this feedback to help revise the set of items to be tested in the structural phase. Engaging in content alignment provides assurance that items map to the proposed dimensions of the scale. In addition, a content alignment provides evidence that items are both relevant to and representative of the full construct (American Educational Research Association et al., 2014).

### 4.2.1. Participants and procedures

Eight judges participated, all of whom had expertise in measurement and motivation theory. Five of the raters held a Ph.D. in educational psychology or a related field, while the remaining three were advanced doctoral students. Experts were provided with an electronic form, where each of the cost dimensions and definitions from Fig. 2 were listed with space to indicate whether they believed each item corresponded to one, more than one, or none of the cost dimensions. A high degree of agreement among the

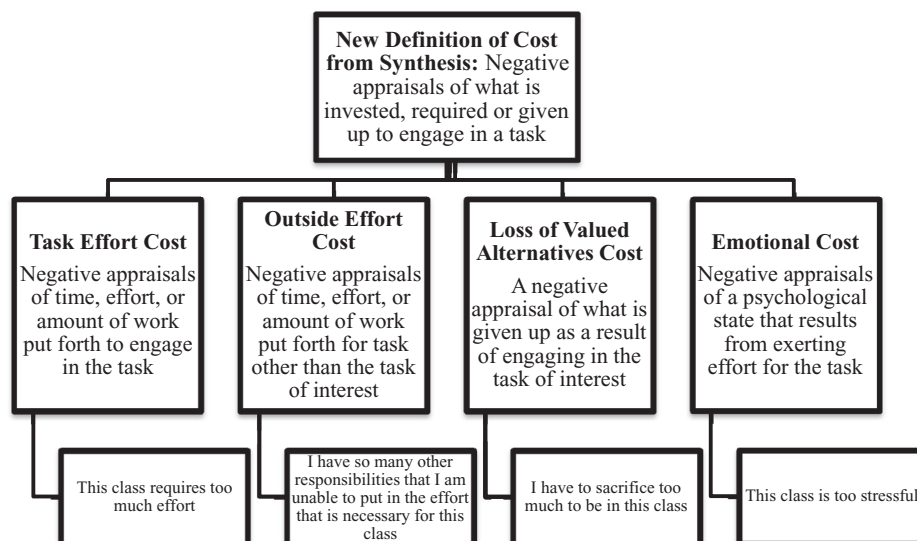


Fig. 2. Operational definitions for each cost dimension.

item reviewers would provide evidence for content validity. Experts were also asked to rate how certain they were about their mapping and how relevant they thought the item was to the dimension on a 1–4 scale, with 1 being, “Very Uncertain” and “Very Irrelevant” and 4 being, “Very Certain” and “Very Relevant.” Finally, experts were provided space to include feedback about the each item. These comments were also considered when evaluating the items.

In using the content alignment results to make decisions about the scale, all three sources of information were considered. Items were only retained if they had high rater agreement (7 of 8 raters agreed on the intended dimension), high certainty and relevance ratings (majority of raters chose a 3 or 4 on the response scale), and the individual item feedback was not negative.

4.2.2. Results and discussion

All but two items had favorable agreement, with 100% of raters mapping each item to its intended cost dimension. However, raters were allowed to map items to more than one dimension and there were two raters who consistently dual-mapped items. Twenty-three items had one or two of these raters map it to more than one dimension. The items and comments regarding those items were carefully considered and upon further review, the items appeared to be measuring two constructs. For example, the items “There is so much work in this class that it causes me to feel stressed out” and “The requirements for this class stress me out” both measure effort and emotional costs, thus they were removed.

In regard to the certainty and relevance, 32 of the items had ratings with the majority of raters choosing “Very Certain” and “Very Relevant.” Items with consistent low ratings were removed or revised. For example the items, “Taking this class makes me unhappy” and “I have too much going on in my life to put time into this task” were removed for low certainty and relevance ratings.

Finally, the experts provided useful feedback about items. Eighteen items were considered confusing, multidimensional, or irrelevant by numerous experts. Those items were removed. Additionally, three items under the Loss of Valued Alternatives subscale were revised to include “too much” language to ensure they conveyed a negative appraisal. The final scale, for use in the quantitative analyses, included 24 items, 7 of which were revised from their original wording.

5. Structural phase

In the structural phase, the emphasis shifts from substantive understanding to psychometric properties. The primary purpose of this phase is to gather evidence that the items used to measure the construct are inter-related in expected ways. We collected two samples, one for an exploratory factor analysis and one for a confirmatory factor analysis, following recommendations made by [McCoach, Gable, and Madura \(2013\)](#). These two samples were used to replicate the hypothesized scale structure.

5.1. Exploratory factor analysis

We began our psychometric investigation with an exploratory factor analysis (EFA) because we had no prior empirical evidence that the theoretical structure we proposed would manifest as four factors. In particular, we were piloting new items for a new dimension of cost that involved outside task effort, and two experts in the content validation indicated they thought some items mapped to numerous dimensions. An exploratory analysis provides an avenue for uncovering factor solutions researchers may not foresee, as well as unanticipated item cross loading. We used EFA results to form competing models that could be formally tested in a confirmatory framework.

5.1.1. EFA participants and procedures

Students from a large, northeastern, public university were recruited from two spring semester, sections of intermediate calculus. Students completed the cost measure with regard to their current calculus class to inform a pilot study investigating motivation changes in calculus, where a high number of students fail, drop, or withdraw from the course. Participation was voluntary and was not tied to course credit. The items were presented in random order via a scantron form, with responses spanning a 9-point scale with 1 = “Completely Agree”, 5 = “Neither Agree nor Disagree”, and 9 = “Completely Agree.” The sample included 184 students, 43.5% male (15% did not indicate gender) with an average age of 19.5 (*SD* = .96, 12% did not indicate age). Participants identified themselves as 66.5% White, 14.1% Asian, 5.2% Hispanic or Latino, 4.7% Black or African American, and 2.6% Multi-ethnic (6.5% did not indicate ethnicity or chose other).

5.1.2. Exploratory factor analyses and results

We conducted the exploratory factor analyses in two parts: we used empirical extraction techniques to determine the optimal number of factors, and then we conducted the EFAs using those results. We compared our hypothesized four-factor solution to results from a Parallel Analysis (cf., [O’Connor, 2000](#)), the Minimum Average Partial Procedure (cf., [O’Connor, 2000](#)), Scree Plot, and the number of eigenvalues greater than 1, giving the greatest weight to the Parallel Analysis and Minimum Average Partial Procedure ([Crawford & Koopman, 1979](#); [Zwick & Velicer, 1986](#)).

[Table 5](#) includes a summary of each extraction technique and the number of factors suggested. The various extraction techniques suggested that 2–4 factors were present in the data. To determine the viability of each solution we extracted each in an EFA using principal axis factoring and direct oblimin rotation. The two factor solution was composed of one factor including the task effort cost, loss of valued alternatives, and emotional cost items. The second factor included the outside effort cost items. The three factor solution included one factor for task effort cost and loss of valued alternatives, a second factor for emotional costs and a third factor for outside-costs. The four factor solution included factors that mapped on to the hypothesized structure, explaining 76% of the item variance. Due to space limitations we only present the pattern matrix from the hypothesized, four factor solution, as this solution seemed

**Table 5**  
Factor extraction techniques and results.

Method	Method description	Number of factors extracted
Eigenvalues <1	Factors correspond to the number of factors that have eigenvalues greater than 1.	3
Number of factors above scree plot elbow	Factors correspond to the area above the elbow of the scree plot.	3
PAF parallel analysis 95% cutoff	Our data are compared to simulated data in which there are no factors. Factors correspond to the ones in which the sample eigenvalues are higher than the random data eigenvalues when compared to a 95% cutoff.	2
Minimum average partial procedure squared	One factor is extracted, and then a partial correlation matrix, taking into account that factor, is computed. This is continued until the average of those squared partial correlations reaches a minimum (variance explained by the additional factors is not increasing).	4
Minimum average partial procedure to the 4th	Similar to squared procedure above, however, the average of the partial correlations to the fourth power is analyzed.	3

**Table 6**Pattern<sup>a</sup> matrix from hypothesized EFA solution with intended factors.

Item <sup>b</sup>	Task effort cost	LOVA <sup>c</sup>	Emotional cost	Outside effort cost
1. This class demands too much of my time.	.805	.149		
2. This class is too demanding.	.796		.119	
3. I have to put too much energy into this class.	.727	–.113		.230
4. This class is too much work.	.916			
5. This class takes up too much time.	.774	.120		
6. This class takes too much effort.	.750		.201	
7. I have to sacrifice too much to be in this class.	.594			.199
8. This class requires me to give up too many other activities I value. <sup>b</sup>	.419			.483
9. This class makes me miss out on other things I care about.	.110	.273	.141	.598
10. Taking this class causes me to miss out on other things I care about.		.267	.237	.590
11. I can't spend as much doing the other things that I would like because I am taking this class.	.190	.259	.129	.434
12. This class is mentally exhausting. <sup>b</sup>	.544		.259	
13. I feel too anxious about this class.	–.123		.874	
14. This class is emotionally draining.	.280		.524	.131
15. I worry too much about this class. <sup>b</sup>	.368		.399	.123
16. This class makes me feel too anxious.			.812	
17. This class is too stressful.	.277		.663	
18. This class takes too much out of me emotionally.			.744	.158
19. This class is too frustrating. <sup>b</sup>	.425	.151	.436	–.182
20. Because of all of the other demands on my time, I don't have enough time for this class.		.893		
21. I have so many other responsibilities that I am unable to put in the effort that is necessary for this class.		.912		
22. I have so many other commitments that I can't put forth the effort needed for this class.	.145	.841		
23. Because of the other things I do, I don't have time for this class.		.949		
24. I can't put the time that I need into this class because of all of my other demands.		.896		

<sup>a</sup> Coefficients are interpreted as the relationship each item has to the factor, after controlling for the other factors.<sup>b</sup> Indicates items with substantial crossloading or poor loading on intended factor.<sup>c</sup> LOVA = Loss of Valued Alternatives.

most tenable. Per recommendations from McCoach et al. (2013), we chose this solution because it had at least three items on each factor with strong primary loadings (greater than .40), which suggests that the solution is not over-extracted. Further, with the exception of one item, all items had strong primary loadings on their hypothesized factor. Only three items showed substantial cross loading (as recommended by McCoach et al., 2013), with loadings greater than .30 on a second factor. Table 6 includes all items and indicates which items had cross loadings.

## 5.2. Confirmatory factor analysis

Though the EFAs provide initial evidence that the hypothesized factor solution may be tenable, it also shows evidence of other, potential solutions and item cross loading. Thus, we retained all items for further investigation. Confirmatory factor analysis (CFA) is needed to formally test the competing solutions and also provides an opportunity to further investigate the potentially problematic items. Therefore, we collected a second sample, keeping all items in their original form.

### 5.2.1. CFA participants and procedures

We collected data from our second sample at the same university and in the same manner as the EFA sample. Again, these data are from a larger study, serving as the first time point in a longitudinal design. This sample was collected during the following fall semester in two sections of introductory calculus. The sample included 228 students; 57% male (13% did not indicate gender) with an average age of 18.7 ( $SD = 4.9$ , 15% did not indicate age). Participants identified themselves as 67.8% White, 13.9% Asian, 3.9% Hispanic or Latino, 1.7% Black or African American, and 2.6% Multi-ethnic (10.1% did not indicate ethnicity or chose other).

Using the results from the factor extraction methods from the EFA as guidance we conducted numerous confirmatory factor analyses in MPlus version 7.1. Table 6 shows all items used in the CFA and their intended factors. We ran four models, collapsing items across their intended factors in various ways. Model A was a

single-factor model with all items as indicators. Model B was a two-factor model with one factor representing task effort cost, loss of valued alternatives and emotional cost and the second factor representing outside effort cost. Model C combined task effort cost and loss of valued alternatives into one factor, emotional cost items into another factor, and then outside effort cost into a third factor. Finally, Model D treated each hypothesized dimension as a separate latent factor. Models were judged using incremental and absolute measures of fit:  $\chi^2$  test of model fit, root mean square error of approximation (RMSEA), the comparative fit index (CFI), Tucker–Lewis index (TLI), and the standardized root mean square residual (SRMR). Further, chi-square differences tests were used to judge differences in model fit between the competing, nested models.

### 5.2.2. Results and discussions

The hypothesized, four-factor model fit significantly better than the competing models that were suggested by the previous exploratory factor analyses (see Table 7). The fit indices for the four-factor model were acceptable, but exhibited room for improvement. Further, results from the EFAs suggested three items with cross-loading, thus we investigated the modification indices and residual correlation matrices. This revealed some problematic items. Specifically, some pairs of items were extremely related, more so than the factor would predict, as they had high modification indices for residual correlations. Upon investigation these item pairs had extremely similar wording. For example, items 9 and 10 from Table 6 both used the phrase “miss out” and items 20 and 24 both use the term “demands.” We removed one of each such redundant items, removing a total of five items. Fit for the trimmed, four-factor model improved substantially, with all indices above acceptable cutoffs (Table 7).

Though the model fit well, we considered the latent factor correlations (Table 8) for each of the four factors. The correlations were very high, suggesting a higher order factor might best represent the data. Thus we tested a four-factor solution against a higher order factor solution in which each of the four lower-order cost factors became an indicator of a higher-order factor, representing a general



**Table 7**  
Fit statistics for the cost scale measurement models from study 5.

Model	$\chi^2$	df	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$	$\Delta df$
Model D: Four-factor model	742.076	246	0.919	0.909	0.094	0.042	–	–
Model C: Three-factor model	851.888	249	0.901	0.890	0.103	0.038	109.812*	3
Model B: Two-factor model	1060.546	251	0.867	0.867	0.119	0.045	318.470*	5
Model A: One-factor model	1333.671	252	0.823	0.806	0.137	0.050	591.595*	6
Trimmed four-factor model	344.522	146	0.956	0.948	0.077	0.029	–	–
Trimmed higher-order four-factor model	357.952	148	0.953	0.946	0.079	0.042	13.430*	2

\* Indicates  $p < .05$ .

**Table 8**  
Correlations between four latent cost sub-factors.

Sub-factor	1	2	3	4
1. Task effort	–			
2. Outside effort	0.849*	–		
3. Loss of valued alternatives	0.946*	0.872*	–	
4. Emotional	0.930*	0.832*	0.868*	–

\* Indicates  $p < .01$ .

cost dimension (see Fig. 3). Model identification was achieved by fixing the higher-order factor variance to one and freely estimating each lower-order path. Though the chi-square difference test between these two models was significantly different (see Table 7), favoring the four-factor model, the fit indices were not practically different and within acceptable range, suggesting the higher-order factor model is appropriate for the data. Further, all lower-order factors and the higher-order cost factor exhibited excellent internal consistency with coefficient alphas above .89. Table 9 includes the final, recommended scale items and reliabilities for each factor.

6. External phase

Though the substantive development and internal structure of the cost scale provides a critical foundation, what educational researchers are ultimately most interested in is how cost relates to other constructs. In the external phase, the importance of the construct is expressed through its relationships with other variables (Benson 1998; American Educational Research Association et al., 2014). We present initial external validity evidence for the cost scale and provide discussion of future areas of research.

6.1. Participants and procedures

To provide evidence of relationships with other, theoretically related variables, we used a subsample of the CFA participants. One of the two sections of fall calculus participants also provided responses to a follow-up survey during their last week of class. Students filled out the cost scale, as well as a brief measure of expectancy and value early in the semester (week 5), and then consented to the release of their final course grade. During the final week of classes (week 15) they filled out measures of continued interest, overall motivation, and reported on their college major. The

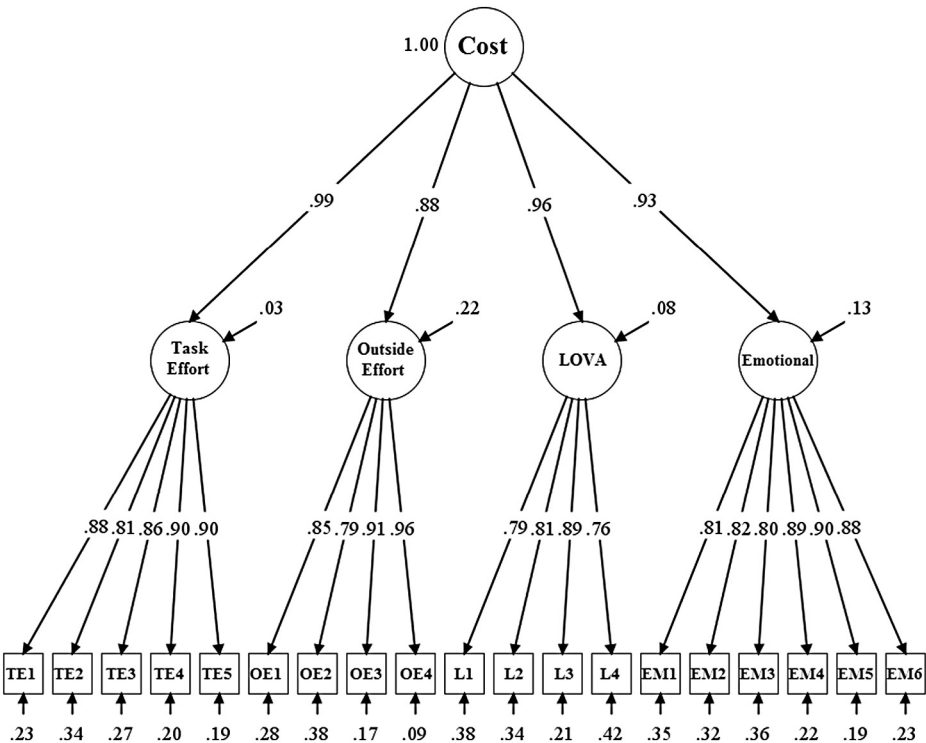


Fig. 3. Final, recommended higher order factor model with standardized coefficients.

**Table 9**

Final cost scale with lower-order factor reliabilities.

<b>Full cost scale: 19 items, alpha = .97</b>	
<b>Task effort cost: 5 items, alpha = .95</b>	
TE1	This class demands too much of my time.
TE2	I have to put too much energy into this class.
TE3	This class takes up too much time.
TE4	This class is too much work.
TE5	This class requires too much effort
<b>Outside effort cost: 4 items, alpha = .93</b>	
OE1	I have so many other commitments that I can't put forth the effort needed for this class.
OE2	Because of the all the other demands on my time, I don't have enough time for this class.
OE3	I have so many other responsibilities that I am unable to put in the effort that is necessary for this class.
OE4	Because of other things that I do, I don't have time to put into this class.
<b>Loss of valued alternatives: 4 items, alpha = .89</b>	
L1	I have to sacrifice too much to be in this class.
L2	This class requires me to give up too many other activities I value.
L3	Taking this class causes me to miss out on too many other things I care about.
L4	I can't spend as much time doing the other things that I would like because I am taking this class.
<b>Emotional cost: 6 items, alpha = .94</b>	
EM1	I worry too much about this class.
EM2	This class is too exhausting.
EM3	This class is emotionally draining.
EM4	This class is too frustrating.
EM5	This class is too stressful.
EM6	This class makes me feel too anxious

Note: Students used a 9-point response scale with 1 = "Completely Agree", 5 = "Neither Agree nor Disagree", and 9 = "Completely Agree".

subsample used for the correlational analyses included 95 students. Participants were 57% male (8% did not indicate gender), identified themselves as 76.8% White, 12.6% Asian, 4.2% Hispanic or Latino, 3.2% Black or African American, and 2.1% Multi-ethnic (1% did not indicate ethnicity or chose other), and had an average age of 18.4 ( $SD = 0.78$ , 11% did not indicate age). Almost all students indicated they had declared a major (93%). When reporting their major, students selected from a discrete number of categories: Sciences (e.g., Life Sciences/Biology, Chemistry, Physics), Applied Sciences (e.g., Nursing, Agriculture, Pharmacy, Allied Health), Engineering/Computer Science, Math/Statistics, Social Sciences, Arts/Humanities, and all other major categories. A majority of the subsample (71%) indicated they were Engineering/Computer Science majors, 13% of the sample reported Sciences, and 8% of the sample reported to be Social Sciences majors, with only a few students reporting the other categories.

**Table 10**

Correlations among cost, expectancy, value, and student outcomes from external phase.

Measure	1	2	3	4	5	6	7	8	9	10
1. Cost-ALL	–									
2. LOVA <sup>a</sup>	.930*	–								
3. Task effort cost	.954*	.883*	–							
4. Outside effort cost	.887*	.767*	.776*	–						
5. Emotional cost	.961*	.852*	.894*	.806*	–					
6. Expectancy	–.554*	–.433*	–.426*	–.602*	–.596*	–				
7. Value	–.366*	–.370*	–.379*	–.284*	–.333*	.151	–			
8. Final grade	–.377*	–.346*	–.301*	–.344*	–.404*	.365*	.029	–		
9. Long-term interest	–.306*	–.293*	–.335*	–.183	–.313*	.117	.599*	.286*	–	
10. Overall motivation	–.373*	–.334*	–.371*	–.373*	–.335*	.289*	.320*	.361*	.510*	–
Mean	3.50	3.30	3.61	3.30	3.68	7.16	6.56	83.96	6.17	6.52
SD	1.53	1.57	1.61	1.64	1.68	1.29	1.31	11.16	1.76	1.58

Note:  $N = 95$  for all analyses.

\* Indicates  $p < .01$ .

<sup>a</sup> LOVA = Loss of Valued Alternatives.

## 6.2. Measures

Students responded to all measures using scantron forms, during class. Items were presented in random order, all items had the same response scale: 9-points with 1 = "Completely Agree", 5 = "Neither Agree nor Disagree", and 9 = "Completely Agree."

### 6.2.1. Cost

The final cost scale (Table 9) was used to create an observed score mean of all 19 items. Then the items comprising each dimension of cost (task effort cost, outside effort cost, loss of valued alternatives and emotional cost) were used to create an observed subscale score mean.

### 6.2.2. Expectancy and value

Four items were used to measure students' overall expectancy for their class. These items captured general expectations for success in the class (e.g., I think I can do well in this class) and provided a reliable mean score (coefficient  $\alpha = .88$ ). Five items were used to measure students' overall value for their class (e.g., I think this class will be useful to me, I think this class is interesting) and provided a reliable mean score (coefficient  $\alpha = .85$ ). These items were adapted from Authors (in press).

### 6.2.3. Outcomes: continued interest, overall motivation and performance

Continued interest is defined as a desire to re-engage in a behavior in the future (Harackiewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008) and was measured with five items (e.g., I am interested in learning more about math, I am likely to go into a career that involves math), which provided a reliable mean score (Cronbach's  $\alpha = .90$ ). Additionally a one-item measure of overall motivation at the end of the course was used in the correlational analyses, "Overall, I am motivated in this class." As a measure of student performance, we used students' final grades in the class on a 0–100% scale.

## 6.3. Results and discussion

Table 10 displays the zero-order correlations for the cost general scale, each cost dimension, expectancy, value, the student's final course grade, continued interest, and overall motivation. As expected from the CFA, the cost subscales were highly correlated with one another and to the general cost factor. Each measure of cost was strongly and negatively related to expectancy and moderately and negatively related to value. Cost was also moderately, negatively correlated to all student outcomes included in the analysis. Correlations

between the cost subscales and outcomes were similar to one another, approximately  $r = -.30$ . As is usually seen in other expectancy-value research (for a review, see [Wigfield & Cambria, 2010](#)), expectancy was strongly correlated to student performance and value to continued interest.

Though preliminary, these results suggest that cost, and the specific dimensions, are related to the other components in the expectancy-value model, as well as important student outcomes. Worth noting is the strong relationship between expectancy and cost, despite the historical view of cost as a value component. These results provide some evidence that cost is tied to both expectancy and value components. We hypothesize that low expectancy causes higher cost, such as stress and anxiety, which was discussed as a fear of failure in Eccles's seminal work. Finally, cost components are significantly related to both performance and long-term interest measures, whereas expectancy is only related to performance and value only related to long-term interest. However, we observed that outside effort cost was the least related to long-term interest. This may be because outside costs can change from semester to semester (i.e., death in the family or financial hardship) for some students, having a relatively weak impact on long-term goals. These results offer promise of the cost construct in better understanding these complex situation factors that influence student behavior and motivation.

## 7. General discussion and future directions

Despite strong recommendations to carefully consider the substantive phase of construct validation ([Borsboom & Mellenbergh, 2004](#); [Gehlbach & Brinkworth, 2011](#); [Gorin, 2007](#); [Zumbo & Shear, 2011](#)), the majority of published validity work focuses on factor analysis and correlational studies (i.e., the structural and external phases of Benson's framework). To correct this shortcoming, we embarked on a series of studies to better understand the substance of cost before moving onto the structural and external phases. Using this strong substantive foundation, we developed operational definitions of cost that incorporated both theory and empirical findings. The result is a set of principles for operationalizing cost and a new scale that can be utilized by the research and practitioner community.

### 7.1. Defining and operationalizing the substance of cost

The substantive phase provided evidence of confusion in the field about how to define and measure cost. Our substantive work also yielded three major extensions of past research. First, we extended the number of dimensions that have been explored in the literature from three to four. In particular, outside effort cost was a new addition to the dimensions that are typically discussed in past work: effort cost, loss of valued alternatives cost, and emotional cost. This measure of outside effort cost is an attempt to capture the costs students experience in other areas of their life, outside of a particular task or class. This outside effort cost may negatively impact their ability to engage and succeed in a particular class (such as working a part-time job to help or responding to a family emergency).

Second, we created an operational definition of cost to distinguish it from measures of general effort, difficulty, or value that may not reflect cost. As is shown in [Tables 1–3](#) and across numerous studies, cost has been measured differently, with different items and substantive focus in each study. As an example, Eccles's earlier work operationalized cost as a value, with items capturing effort and value (e.g., Is the amount of effort it will take to do well in your math course this year worthwhile to you?), whereas [Luttrell et al. \(2010\)](#) operationalized cost as difficulty (e.g., I have to study much harder for math than for other courses). By creating a cost definition that is contingent on a negative appraisal (e.g., This class is too much work), the cost item is distinguishable from general effort, difficulty,

or value. The “too much” language in the items captures a pivotal difference in the student's subjective experience. This approach is further bolstered by the results of our qualitative study. When reflecting on their most motivating class, students relish in the hard work, but when reflecting on their least motivating class they appraise their hard work negatively, as too much. This negative appraisal can help us to distinguish the motivated student from the unmotivated one.

Literature outside of educational psychology provides support for the idea that an objective cost (such as amount of effort needed) can be perceived subjectively in one of two ways. The Job Demands and Resources (JD-R) model ([Demerouti, Bakker, Nachreiner, & Schaufeli, 2001](#)), a model from I/O psychology, has a component called job demands that resembles Eccles's conceptualization of cost, but studies using the JD-R model have found some job demands to be positive in nature ([Mauno, Kinnunen, & Ruokolainen, 2007](#)) and related to employee engagement. To address the contradiction in the JD-R model, [Van den Broeck et al. \(2010\)](#) split job demands into two parts: job challenges and job hindrances and investigated their relationship with employee vigor. Even though job challenges and job hindrances both involve energy depleting costs, only job hindrances were negatively related to vigor, whereas job challenges were positively related to vigor.

Finally, we worked to create a scale that could be used in numerous contexts. We were careful to write items that were not specific to a certain subpopulation or academic context (e.g., college students or graduate students). This is a different approach than those taken by previous researchers (see [Tables 2 and 3](#)). We recognize that the costs experienced will vary from student to student. For example if loss of valued alternatives is high, we would expect students to agree to the statement, “Taking this class forces me to give up too many other things I care about”, but they may not agree with, “Taking this class forces me to give up time with family” if they do value time with family. Such an approach allowed us to create a measure that would be more appropriate for a wider range of students and academic tasks, which will facilitate the generalizability of future cost research. However, because of the general nature of the scale, it is limited in identifying specific costs students may experience. Understanding why cost occurs is an important area of future research that may be better understood through qualitative methodology, such as the work conducted by [Watkinson et al. \(2005\)](#). Writing quantitative items to cover the particular costs of all subpopulations or academic contexts may not be feasible or practical.

### 7.2. Is cost uni-dimensional or multi-dimensional?

The studies conducted for the structural phase provided the strongest support for the 4-factor solution, where each dimension of cost is treated as a separate factor. This solution is more nuanced than cost measurement seen in some previous work ([Luttrell et al., 2010](#); [Trautwein et al., 2012](#)), and is more akin to the sub-factors used in [Perez et al. \(2014\)](#). However, the higher order factor model, incorporating a general cost factor, also fit the data well. The question of how to model cost, as a higher-order factor, or four-highly correlated dimensions, warrants further study and may depend on the research question.

In the current sample, the scale showed adequate reliability and model fit as a four factor scale, or as a four factor scale with a higher-order, general cost factor. Our research is limited in that we only collected data from calculus students in a particular academic context, and for these students the cost dimensions were highly correlated. A critical area of future research is to investigate the empirical structure of cost under different circumstances and with different groups of students to determine whether our findings replicate across other types of students. This type of research is crucial for understanding

how the scale functions, but more importantly it provides substantive insights as to how cost manifests differentially in different situations.

Though our four factors were highly related, we did see differential relationships in our correlational study. We observed that outside effort cost was more related to expectancy than any other component. These findings beg the question of how demands outside of class influence expectancy and the causal relationship between the two constructs. Further, emotional cost was more related to final grade than any other construct or cost component. Emotional stress may be more predictive of performance whereas other types of costs may be more predictive of outcomes we were unable to include in our study, such as persistence in a subject.

### 7.3. Limitation and areas for future research

Despite our efforts to be thorough in our development of a cost scale, this work is not without limitations and there is more research to be done. First, there were some limitations in our qualitative study. We asked students to reflect on their classes. Those retrospective appraisals could be different than real time appraisals. Thus, more qualitative research is needed to better understand differences in current versus retroactive cost.

Second, we found an emerging theme: *not enough cost*. Students described, when least motivated, that the class did not require enough effort, or was “too easy.” We chose not to incorporate this into our current cost scale because we believe this phenomenon is captured by Csikszentmihalyi’s (1975) Flow Theory. This theory of motivation proposes that optimal motivation is a function of skill and challenge. If challenge is low, but skill is high, one experiences boredom. Conversely, one experiences anxiety when challenge is too high for the skill level. Perhaps, when *not enough cost* is experienced, students are experiencing a negative psychological cost of boredom or a cost of loss of valued alternatives. However, more research is needed to understand how a student’s state of flow influences their perceived costs.

Third, we only presented initial evidence from the external phase of construct validation. For example, we found cost subscales to be negatively correlated to general motivation, expectancy, value, performance, and continued interest. The subscales demonstrated this property with moderate to strong negative correlations. However, this is a crude validity check at best and more external studies are needed to understand how cost fits into the larger theoretical realm of motivation theory. Interestingly, we observed that cost components were moderately related to both continued interest and performance, whereas expectancy was only related to performance and value only related to continued interest. Cost seems to relate to these outcomes in ways that expectancy and value do not, but more study is needed. As we move forward with cost research, investigating the differential predictability of expectancy, value, and the distinct cost components is imperative.

Finally, as discussed before, we only sampled college students in our qualitative study and college, calculus students in our psychometric investigation. In addition, the majority of calculus students who participated reported that they were pursuing a major in a STEM field (i.e., science, technology, engineering, or math). As discussed in Perez et al. (2014), STEM students are unique in their motivational attributes. The structure of the scale and relationships to outcomes could be different with different samples of students. An important area of future research is investigating the invariance of our scale with students from other majors and other age groups, as their attitudes could be formed differently. Although STEM students may be particularly likely to report and experience cost due to the high rates of dropout and attrition from STEM disciplines, feelings of cost certainly are not specific to STEM students. Any educational environment could create cost. In fact, recent pushes to

increase rigor in K-12 and higher education should by definition increase the overall effort that is required of students, which could manifest as a cost.

In the beginning of this article, we posed four questions about students’ descriptions of cost: What is it that these students are describing? Can it be measured systematically? How is it related to students’ motivation and academic performance? And, what could teachers do to optimize student motivation if they knew students were experiencing it? The current set of studies offers answers to the first three questions that support and extend the emerging body of research on cost. The last question, “What could teachers do to optimize student motivation if they knew students were experiencing it?” remains unanswered and is an exciting area for future research. However, an important first step for the field is developing valid and reliable tools to measure and capture cost.

### Acknowledgments

Part of this work was presented as a Master’s Thesis at James Madison University. The authors would like to thank Eric Hehman, Dena Pastor, Amit Savkar, and Kelli Samonte for their theoretical insights, feedback, and support.

### References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Joint Committee on Standards for Educational and Psychological Testing.
- Anderson, R. D., & Thelk, A. D. (2005). Community college strategies in instrument selection. *Assessment Update*, 17(2), 14–16.
- Barron, K. E., & Hulleman, C. S. (2015). The expectancy-value-cost model of motivation. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed.). Oxford: Elsevier Ltd.
- Battle, A., & Wigfield, A. (2003). College women’s value orientations toward family, career, and graduate school. *Journal of Vocational Behavior*, 62(1), 56–75. doi:10.1016/S0001-8791(02)00037-4.
- Benson, J. (1998). Developing a Strong Program of Construct Validation: A Test Anxiety Example. *Educational Measurement: Issues and Practice*, 17(1), 10–17.
- Borsboom, D., & Mellenbergh, G. J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061–1071. doi:10.1037/0033-295X.111.4.1061.
- Chen, A., & Liu, X. (2009). Task values, cost, and choice decisions in college physical education. *Journal of Teaching in Physical Education*, 28, 192–213.
- Chiang, E. S., Byrd, S. P., & Molin, A. J. (2011). Children’s perceived cost for exercise: Application of an expectancy-value paradigm. *Health Education & Behavior: The Official Publication of the Society for Public Health Education*, 38(2), 143–149. doi:10.1177/1090198110376350.
- Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology*, 104(1), 32–47. doi:10.1037/a0026042.
- Crawford, C. B., & Koopman, P. (1979). Note: Inter-rater reliability of scree test and mean square ratio test of number of factors. *Perceptual and Motor Skills*, 49, 223–226.
- Creswell, J. W., Hanson, W. E., Clark Plano, V. L., & Morales, A. (2007). Qualitative research designs: Selection and implementation. *The Counseling Psychologist*, 35(2), 236–264. doi:10.1177/0011000006287390.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco: Jossey Bass.
- Davis, R. V. (1987). Scale construction. *Journal of Counseling Psychology*, 34(4), 481–489. doi:10.1037//0022-0167.34.4.481.
- Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, 86(3), 499–512. doi:10.1037//0021-9010.86.3.499.
- Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A. S. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105–121). New York: The Guildford Press.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–138). San Francisco: W.H. Freeman and Company.
- Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents’ achievement talk values and expectancy related beliefs. *Personality and Social Psychology Bulletin*, 21(3).
- Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology*, 15(4), 380–387. doi:10.1037/a0025704.
- Gorin, J. S. (2007). Test design with cognition in mind. *Educational Measurement: Issues and Practice*, 25(4), 21–35. doi:10.1111/j.1745-3992.2006.00076.x.



- Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology*, 100(1), 105–122. doi:10.1037/0022-0663.100.1.105.
- Hesse-Biber, S. N., & Leavy, P. (2011). *The practice of qualitative research* (2nd ed.). Thousand Oaks, CA: SAGE Publications, Inc.
- Luttrell, V. R., Callen, B. W., Allen, C. S., Wood, M. D., Deeds, D. G., & Richard, D. C. S. (2010). The mathematics value inventory for general education students: Development and initial validation. *Educational and Psychological Measurement*, 70(1), 142–160. doi:10.1177/0013164409344526.
- Madden, G. J. (2000). A behavioral economics primer. In W. K. Bickel & R. E. Vuchinich (Eds.), *Reframing health behavior change with behavioral economics* (pp. 13–22). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mauno, S., Kinnunen, U., & Ruokolainen, M. (2007). Job demands and resources as antecedents of work engagement: A longitudinal study. *Journal of Vocational Behavior*, 70(1), 149–171. doi:10.1016/j.jvb.2006.09.002.
- McCoach, D. B., Gable, R. K., & Madura, P. J. (2013). *Instrument development in the affective domain: School and corporate applications* (p. 300). New York: Springer.
- Messick, S. (1980). Test validity and the ethics of assessment. *American psychologist*, 35(11).
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32(3), 396–402. <<http://www.ncbi.nlm.nih.gov/pubmed/11029811>>.
- Parsons, J. S., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1980). *Self-perceptions, task perceptions and academic choice: origins and change* (pp. 1–98). Ann Arbor, MI. Retrieved from: <<http://www.rcgd.isr.umich.edu/garp/articles/parsons80b.pdf>>.
- Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology*, 106(1), 315–329. doi:10.1037/a0034027.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667–686. doi:10.1037/0022-0663.95.4.667.
- Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47(2), 149–155.
- Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology*, 104(3), 763–777. doi:10.1037/a0027470.
- Van den Broeck, A., De Cuyper, N., De Witte, H., & Vansteenkiste, M. (2010). Not all job demands are equal: Differentiating job hindrances and job challenges in the Job Demands–Resources model. *European Journal of Work and Organizational Psychology*, 19(6), 735–759. doi:10.1080/13594320903223839.
- Watkinson, E. J., Dwyer, S. A., & Nielsen, A. B. (2005). Children theorize about reasons for recess engagement: Does expectancy-value theory apply? *Adapted Physical Activity Quarterly*, 22, 179–197.
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, 30(1), 1–35. doi:10.1016/j.dr.2009.12.001.
- Zumbo, B., & Shear, B. (2011). The concept of validity and some novel validation methods. In Presentation made at the Annual Meeting of the Northeastern Education Research Association, Rocky Hill, CT.
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432–442. doi:10.1037//0033-2909.99.3.432.