Assessing Writing 48 (2021) 100509

Contents lists available at ScienceDirect

Assessing Writing

journal homepage: www.elsevier.com/locate/asw

Writing motivation: A validation study of self-judgment and performance

Guangming Ling [a], Norbert Elliot [b,*], Jill C. Burstein [a], Daniel F. McCaffrey [a], Charles A. MacArthur [c], Steven Holtzman [a]

[a] *Educational Testing Service, 660 Rosedale Rd, Princeton, NJ, 08541, United States*
[b] *New Jersey Institute of Technology, University Heights, Newark, NJ, 07102, United States*
[c] *University of Delaware, 106 Alison Hall West, Newark, DE, 19716, United States*

ARTICLE INFO

Keywords:
Writing assessment
Writing motivation
Academic performance
Automated writing evaluation (AWE)

ABSTRACT

This study reports on validation of a writing motivation survey and its relationship with a variety of indicators of academic performance of 566 undergraduate students drawn from six US post-secondary institutions. A writing motivation survey was used to capture students' writing goals, confidence, beliefs, and affect. Two research questions are addressed in the study: 1) What is the internal factor structure of the writing motivation survey completed by a broad range of college students? (2) Using this factor structure, how are the subconstructs of writing motivation associated with the outcomes variables of student performance and features of student writing? Our results confirmed a unidimensional structure for confidence and affect, a 3-factor structure for goals, and a 2-factor structure for beliefs. Low level but significant correlations were also identified between motivation and outcome performance measures, as well as between motivation and features of student writing. The study findings yield insights about the relationship between students' writing motivation, college success indicators, and writing domain knowledge. These findings suggest directions for pedagogical interventions targeting both motivation and writing feature use.

1. Introduction

As East (2019) observes in recalling the history of this journal, writing assessment over the past quarter-century has been concerned with gathering and presenting evidence of validity, reliability, and fairness. In this culture of evidence, as he notes, the construct of motivation has become of increasing significance: Students who are motivated are likely to produce better outcomes. The present study finds its place in a period in which writing motivation is receiving attention by the educational researcher and practitioner communities. As recent attention to sociocultural and sociocognitive views of situated language use continue to advance (Bazerman, 2016; Mislevy, 2018), researchers are presently in an excellent position to take advantage of sophisticated motivational theories (Abdel Latif, 2019; Pintrich, 2000b). In turn, these theories can be related to sophisticated descriptions of the writing construct in which motivation is linked to strategies (Graham, 2018). Programs of research focusing on motivation are needed, especially in terms of the motivational construct's internal structure and its relationship with indicators of college writing performance and other college success indicators.

* Corresponding author.
E-mail addresses: GLing@ets.org (G. Ling), elliot@njit.edu (N. Elliot), jburstein@ets.org (J.C. Burstein), dmccaffrey@ets.org (D.F. McCaffrey), macarthu@udel.edu (C.A. MacArthur), sholtzman@ets.org (S. Holtzman).

https://doi.org/10.1016/j.asw.2020.100509
Received 22 July 2020; Received in revised form 31 October 2020; Accepted 18 November 2020
Available online 1 February 2021
1075-2935/

Equally important, motivation influences student goals for learning and performance. Motivation is an important part of writing instruction because students need to be motivated to achieve, and it is an important outcome of writing instruction as part of continued learning (Bruning & Horn, 2000).

In terms of the assessment of motivation, surveys are commonly used to measure motivation, and evidence of validity and reliability are, in turn, used to draw inferences in terms of further research and pedagogical interventions. Common methods used in existing research are mostly focused on exploratory factor analysis (e.g., Principal Component Analysis), and sampling plans are often limited to a single site (Wright, Hodges, & McTigue, 2019). As well, validity evidence in earlier research has been limited, mainly focusing on one or two criterion variables such as writing score based on a particular course or grade point average (MacArthur, Jennings, & Philippakos, 2019; MacArthur, Traga-Philippakos, May, & Compello, 2019). To expand research on writing motivation, studies are needed using confirmatory factor analysis and structural equation modeling. Specifically, these techniques allow explicit and formal statistical testing of hypotheses in terms of the relationships between observed answers (to survey questions) and related latent constructs (writing motivations) either purely from a theoretical perspective or a mix of theoretical and practical perspectives (Raykov & Marcoulides, 2000). As well, cross-institutional sampling plans can be used to provide evidence of response diversity, and a broadened range of cross-institutional criterion variables can be used to further explore relationships between self-judgements of motivation and academic performance.

To address these needs, in the present study we use a cross-institutional sample to examine criterion-related evidence between writing motivation and a series of external indicators of student success, including standardized admission tests and progression tests, current and future grade point average, and language features of writing samples. Understanding connections between motivation and linguistic features of writing has the potential to improve instruction. In addressing these study aims, ours is the first foundational study to examine relationships among motivation and these related academic measures.

We explore these aims using sampling plan of 566 undergraduate students drawn from six diverse US postsecondary institutions. Following a literature review based on components of the survey and targeted feature analysis using automated writing evaluation (AWE), we use exploratory and confirmatory factor analysis approaches to investigate the internal structure of the survey. Using this factor structure, we then examine relationships among the factors of writing motivation, academic performance (measured through standardized admission tests, concurrent outcomes assessments of writing and critical thinking, concurrent and predictive grade point average), with particular attention to features of student writing samples (as measured by AWE). After identification of study limitations and a brief discussion of our findings, the paper concludes with directions for pedagogical interventions, targeting both motivation and writing feature use, derived from the constructs and subconstructs identified in the study.

2. Literature review

Writing, particularly academic writing, is a challenging task that requires substantial motivation. Flower (1979) viewed composing as a "demanding cognitive operation" (p. 36) in which a goal of writing instruction was giving students "the confidence and motivation to go on" (p. 37). Motivation is influenced both by external social and situational factors and by internal cognitive and affective factors. As such, it has been studied using multiple theoretical frameworks (for reviews, see Abdel Latif, 2019; MacArthur & Graham, 2016).

The most extensive body of research on writing motivation has focused on self-efficacy, defined as individuals' judgments of how well they can accomplish specific tasks (Bandura, 1997). Self-efficacy has been found to predict engagement, persistence, and achievement across a wide range of academic tasks (Bandura, 1997; Pajares, 1996). People tend to engage in activities that make them feel competent. In writing, research has generally found that self-efficacy predicts writing performance even after controlling for prior performance (for reviews, see Bruning & Kauffman, 2016; Pajares & Valiante, 2006). Self-efficacy is most often measured by asking individuals to rate their confidence in completing specific tasks. Several studies have found separate dimensions of writing self-efficacy for composition/organization and for conventions (e.g., Pajares, 2007; Shell, Murphy, & Bruning, 1989; Shell, Colvin, & Bruning, 1995). An early study (Zimmerman & Bandura, 1994) with college students assessed self-efficacy for self-regulation of writing and found positive correlations with grades in first-year composition. One study (Bruning, Dempsey, Kauffman, McKim, & Zumbrunn, 2013) found unique dimensions of self-efficacy for idea generation, conventions, and self-regulation. A recent study (Wright et al., 2019) used a scale that combined self-efficacy and self-concept and found positive correlations with essay quality and a standardized writing assessment.

Writing motivation has also been studied using achievement goal theory (Elliot & Church, 1997; Pintrich, 2000a). The theory includes three contrasting goal orientations toward learning: mastery, performance-approach, and performance-avoidance. Mastery goals are focused on developing knowledge and competence; performance-approach goals refer to attempts to appear competent compared to others; and performance-avoidance goals describe efforts to avoid unfavorable judgments by others. Research on academic motivation has found that mastery goals are related to interest, persistence, strategic learning, and achievement; performance goals are related to achievement; and avoidance goals predict low achievement and interest (for a review, see Senko, Hulleman, & Harackiewicz, 2011). Research on writing goal-orientation has found positive relationships between writing grades and both mastery and performance goals (Kauffman et al., 2010). In addition, three studies (Kauffman et al., 2010; Pajares, Britner, & Valiante, 2000; Pajares & Cheong, 2003) also found that self-efficacy is positively correlated with mastery goals and negatively correlated with avoidance goals. The negative correlations between self-efficacy and avoidance goals are consistent with earlier research on writing apprehension (Daly & Miller, 1975) and writer's block (Rose, 1980) showing negative effects on writing quality; apprehension about writing can lead to avoidance of writing and lower achievement.

Few studies have examined the influence of students' beliefs about what is important to good writing. White and Bruning (2005) developed a scale to assess two sets of beliefs about writing: transformational beliefs that writing is a process of constructing meaning and transmission beliefs that writing is a process of transmitting knowledge of experts; higher achieving college writers scored higher on transaction and lower on transmission beliefs than lower achieving writers. Wright et al. (2019) included a scale of writing beliefs that focused on the perceived value of becoming a good writer; scores of high school students correlated with writing achievement.

In order to advance research in the intrapersonal domain of writing, MacArthur, Philippakos, and Graham (2016) designed a measure of motivation and applied it in a study of basic college writers (n = 131). Their survey measure focused on four subconstructs of motivation—goal orientation, self-efficacy, beliefs, and affect. A Principal Component Analysis approach was used to explore the internal structure of each subscale of writing motivation. Affect was simply defined as liking writing and finding it satisfying; a single dimension was found. As anticipated based on prior research, the factor analysis for goal orientation identified separate factors, or subscales, for mastery, performance, and avoidance using Principal Component Analysis. Based on foundational research showing that basic writers tend to emphasize problems with errors in writing (Shaughnessy, 1977), the beliefs scale found two factors for beliefs in the importance of content and the importance of conventions to good writing. However, although the self-efficacy scale was designed to have three separate factors—writing tasks, writing strategies, and self-regulation—only a unidimensional structure was identified. Reliabilities for all scales ranged from .72 to .95. When analyzed in their relationship to writing performance, two of seven motivational scales were found to have statistically significant but low and negative correlations with achievement measures, ranging from -.27 to -.32 for conventions beliefs, and from -.20 to -.35 for goal avoidance. Comparisons across different levels of basic writing classes found that students in higher-level classes had significantly higher scores for self-efficacy, mastery and performance goals, and lower scores for conventions beliefs and avoidance goals. A pretest to posttest design identified changes in motivation after a semester of instruction that were statistically significant and in the anticipated direction of positive change for six of seven motivational factors.

Subsequently, Philippakos and MacArthur (2015) conducted a follow-up study in which a wider range of students, including first-year composition students and developmental students in four-year and community colleges (n = 371) were included and grammar items were added to the self-efficacy scale. An exploratory factor analysis (Principal Component Analysis) approach was employed. As hypothesized, three factors were found for the self-efficacy, three factors for goal orientation, two for writing beliefs, and one for affect. Writing achievement (quality of a single essay) correlated with all three self-efficacy factors (range .20–.29), but not with any other motivation measures. Although the second study included a somewhat more diverse sample of students (but only developmental level and first-year composition students), it only included a single measure of writing achievement. Further research is needed to evaluate the motivation scale with a larger and more diverse population of college writers, with multiple measures of writing performance and academic achievement, and using confirmatory factor analysis to test the model-data fit.

Writing motivation and writing performance are related to writing success (De La Paz & Graham, 2002). Expressed in more detail, it may reasonably be said that overall student success in writing—the ability to set and sustain goals, navigate writing environments, succeed on diverse genres and topics, and effectively manage composing process, knowledge, and skills (Graham, Gillespie, & McKeown, 2013)—is related to motivation (Bruning & Kauffman, 2016; MacArthur & Graham, 2016; Pajares & Valiante, 2006; Wright et al., 2019). However, in studies of motivation, examining relationships between motivation and performance has been limited because the performance domain is often represented by holistic scores on standardized writing tasks or summative measures such as course grades. Analysis of writing using feature analysis would provide an association with an intrapersonal factor that could be used to describe the roles of performance and motives in writing success. Feature analysis therefore becomes an important, though under investigated, area of research related to motivation and writing performance.

With the advancement of natural language processing (NLP) tools and techniques, analysis of textual features by hand coding has been supplemented by the ability of NLP tools to computationally analyze texts to capture a wide range of writing features, including English conventions (e.g., grammar errors), topic development, mood, coherence (e.g., topic flow, use of transition terms), and argumentation (e.g., use of claims and sources) (Burstein, Riordan, & McCaffrey, 2020; Cahill & Evanini, 2020; Matsumura, Correnti, Zook-Howell, Walsh, & Bickel, 2020). AWE methods have been most commonly used to detect writing construct-relevant features (e.g., grammar errors) which are then modelled using statistical methods (e.g., linear regression) to evaluate quality in student or test-taker writing (Shermis & Burstein, 2013). The technology has been widely used in educational assessment and instruction typically through generating a series of features that capture various aspects of an essay and using these features to either approximate an overall score for the essay (e.g., high-stakes writing assessment) or provide feedback on specific aspects of writing, such as in instructional applications (Cotos, 2011; Ranalli, Link, & Chukharev-Hudilainen, 2017; Warschauer & Ware, 2006; Yannakoudakis, Andersen, Geranpayeh, Briscoe, & Nicholls, 2018). Using the NLP tool Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014), MacArthur, Jennings et al. (2019) and MacArthur, Traga-Philippakos et al. (2019) developed a model of linguistic constructs to predict writing quality for argumentative writing by postsecondary basic writers. AWE has also been shown to have utility in examining relationships between writing features and other outcomes, such as reading skill (Allen, Snow, Crossley, Jackson, & McNamara, 2014; Allen, Dascalu, McNamara, Crossley, & Trausan-Matu, 2016), and academic success factors (Beigman Klebanov, Burstein, Harackiewicz, Priniski, & Mulholland, 2017).

In summary, the literature clearly demonstrates the importance of writing motivation in writing skills and their development and improvement. It is widely acknowledged that a reliable and accurate assessment of writing motivation is also a critical prerequisite in research to uncover factors that may be related to writing motivation, its relationship with writing skills, and other types of outcomes. The significance and prerequisite nature of motivation thus touches on both its internal structure, or construct validity, as well as external validity evidence. Our review suggests that earlier research related to the assessment and validation of writing motivation appears to have the following limitations: (1) findings were based on a relatively small sample with limited variation of writing ability in a single institution; (2) studies relied on an exploratory approach with no confirmatory approach employed; (3) researchers employed a single or limited number of criterion variables of college outcomes and achievement. Further research is therefore needed to evaluate the motivation scale with a larger and more diverse population of college writers, with multiple measures of writing

Table 1
Survey Sample Distribution by Institution[a].

Site	Classification	Enrollment Profile	Student Sample and Population		Gender: M	Gender: F	Race/Ethnicity: American Indian or Alaskan Native	Asian	Black or African American	Hispanic/ Latino	Native Hawaiian or other Pacific Islander	White	Two or more races	Unknown
Institution 1	Master's Colleges & Universities: Larger Programs	Very high undergraduate; Four-year, full-time, selective, lower transfer-in	Population	~9,000	42.1	58.0	0.3	1.1	6.6	7.4	0.0	77.1	1.7	5.8
			Study Sample	18	38.9	61.1	0.0	0.0	0.0	11.1	5.6	83.3	0.0	0.0
Institution 2	Master's Colleges & Universities: Larger Programs; Historically Black University	High undergraduate	Population	~6,000	39.4	60.6	0.2	1.5	82.0	4.5	0.2	2.0	4.2	5.5
			Study Sample	165	33.9	66.1	0.0	0.6	88.5	4.8	0.0	1.2	2.4	2.4
Institution 3	Doctoral/ Professional Universities	High undergraduate	Population	~24,000	41.0	59.0	0.4	12.9	2.6	53.9	0.2	17.8	2.6	9.6
			Study Sample	152	42.1	57.9	0.0	15.8	4.6	63.2	0.0	7.9	2.6	5.9
Institution 4	Master's Colleges & Universities: Larger Programs	Very high undergraduate	Population	~9,000	40.8	59.2	0.3	0.8	19.1	0.5	0.0	68.9	0.0	10.4
			Study Sample	134	39.6	60.4	0.0	0.7	17.9	0.7	0.0	73.1	0.0	7.5
Institution 5	Master's Colleges & Universities: Larger Programs	Professions plus arts & sciences, some graduate coexistence	Population	~9,000	44.3	55.7	0.1	1.0	4.6	2.5	0.1	85.3	4.1	2.4
			Study Sample	51	62.7	37.3	0.0	2.0	3.9	3.9	0.0	84.3	2.0	3.9
Institution 6	Doctoral Universities: High Research Activity	Balanced arts & sciences/professions, some graduate coexistence	Population	~17,000	36.9	63.1	0.3	2.1	4.0	7.5	0.1	77.7	4.0	4.3
			Study Sample	46	34.8	65.2	0.0	0.0	4.3	8.7	0.0	78.3	4.3	4.3
Total			Population	~74,000	40.7	59.3	0.3	3.2	19.8	12.7	0.1	54.8	2.8	6.4
			Study Sample	566	40.3	59.7	0.0	4.8	32.0	20.0	0.2	36.4	1.9	4.8

[a] Classification, enrollment profile, and student population information were obtained online from Carnegie Classification of Institutions of Higher Education. Gender and ethnicity distributions were obtained from National Center for Education Statistics IPEDS College data 2019–2020.

performance and academic achievement, and using confirmatory factor analysis to test the model-data fit and establish or verify an internal structure.

Our motivation for the present study was therefore aimed at addressing these three limitations. As well, we identified a fourth motivation: We sought to identify models for instructional interventions in which motivation and writing performance, understood as malleable factors, could be interwoven in order to improve student academic performance. As we will show below, linking research to pedagogy is especially important as we consider ways that consequential perspectives on writing assessment can be used to support instruction.

3. Research questions

The present study is aimed at answering two specific research questions:

(1) What is the internal factor structure of the writing motivation survey (targeting the subconstructs of writing goals, self-efficacy, beliefs, and affect) completed by a broad range of college students?
(2) Using this factor structure, how are the subconstructs of writing motivation (writing goals, self-efficacy, beliefs, and affect) associated with the outcomes variables of student performance and features of student writing?

4. Survey design

The present study adopted a revised writing motivation survey reported in Philippakos and MacArthur (2015) with deletion of three items from the performance goal scale that had ceiling effects. The questionnaire included four scales (or components): (1) writing goals, with 11 items presented on a 5-point response anchored at response 1 [does not describe me at all], response 3 [somewhat describes me], and response 5 [describes me very well]; (2) writing confidence (self-efficacy), with 22 items presented on a 100 point response anchored at 0% [no chance], 20–30 % [some chance], 50 % [50/50 chance], 70–80 % [good chance], and 100 % [completely sure]; (3) writing beliefs, with 12 items presented on a 5-point response anchored at each point from 1 [strongly disagree] to 5 [strongly agree]; and (4) feelings about writing (affect), with 5 items presented in the same 5-point response as the writing beliefs scale. Items from the scales are shown in Table 2 through Table 5.

5. Sampling plan

Survey data were collected at 6 institutions (see Table 1) as follows: In academic semester Fall 2017, survey data were collected at all 6 institutions; in academic semester Spring 2018, survey data were collected from 3 institutions (Institutions 2, 3, and 4); and, in academic semester Fall 2018, survey data were collected from 1 institution (Institution 4). During that period, there were 16.8 million undergraduates in degree granting institutions in the United States (McFarland et al., 2019). Of these students, 56.2 % were female and 43.8 % were male. In terms of race/ethnicity, 53.0 % of enrolled undergraduates were White, 21.1 % were Hispanic, 7.6 % were Asian and 13 % were Black. Among the sampled schools, institutions 1, 4, 5, and 6 reported notably higher rates of White student enrollment than the national average, whereas Institutions 2 and 3 reported much lower rates. Institution 3 reported a somewhat higher proportion of Hispanic student enrollment than the national average. In all the remaining institutions, the reported rates of Hispanic students were substantially below the national average. Except for Institution 3, each institution in the study reported substantially lower rates of Asian student enrollment. Most of the participating universities had lower rates of Black student enrollment than the nation, except Institution 2 (categorized as a historically Black College and University) and Institution 4. While the present study is the first large-scale attitude study to be conducted at multiple US postsecondary universities and represents a diverse population of undergraduate students, the results should be understood in light of subgroup membership across the sites.

6. Other measures

To gather criterion-related evidence, we administered the HEIghten™ Critical Thinking and Written Communication assessment, two components of a postsecondary outcomes assessment suite. The Critical Thinking assessment was administered in a single testing session. In 45 min, each test taker answers 26 selected response questions. The assessment is designed to address two dimensions of critical thinking: analytic skills (e.g., analyzing and evaluating arguments, evaluating evidence and its use, understanding the language of argumentation, distinguishing between valid and invalid arguments) and synthetic skills (e.g., understanding implications and consequences, developing arguments that are valid and sound; see Liu, Mao, Frankel, & Xu, 2016 for more detailed definitions). The Written Communication assessment is designed to address four dimensions of writing: knowledge of social and rhetorical situations; knowledge of conceptual strategies; knowledge of the writing process; and knowledge of language use and conventions (Rios, Sparks, Zhang, & Liu, 2017). The assessment includes two parts: an essay written to a constructed response argument task; and a section with 24 selected-response questions. The selected-response section has two question sets, each with 12 questions based on a common reading material. The essay is scored for overall quality by one human rater and an AWE program.

Beyond these standardized outcomes measures, we collected existing admissions test scores on mathematics and language achievement, reported individually, as well as total admissions test scores. We also collected grade point averages (GPA) during the semester when students took the survey, along with GPAs in each of the four semesters after that semester. These performance and outcomes indicators are given in Table 7. As noted, the essays completed in the HEIghten Written Communication assessment were

Table 2
Writing Goals Items and Subscales.

When I am writing I am trying to:	Mean	SD	Subscales		
			Mastery	Performance	Avoidance
1. improve how I express my ideas.	4.04	0.90	.64***		
2. keep people from thinking I'm a poor writer.	3.70	1.29		.53***	.23***
3. hide that I have a hard time writing.	2.60	1.26			.78***
4. become a better writer.	4.24	0.87	.66***		
5. have my classmates believe I can write well.	3.13	1.30		.62***	.32***
6. avoid making mistakes in front of my classmates.	3.24	1.32		.56***	.50***
7. persuade others with my writing.	3.77	1.05	.68***		
8. be a better writer than my classmates.	3.02	1.29		.59***	
9. hide how nervous I am about writing.	2.47	1.31			.82***
10. get my teacher to think I am a good writer.	3.81	1.14		.77***	
11. better organize my ideas.	4.19	0.93	.76***		
Mastery					
Performance			.59***		
Avoidance			.01	.27***	
Cronbach's alpha est.			.73	.77	.77

Note that all empty cells at the intersection of item and subscales were set to be zero when fitting the model, meaning related items were not set as related subscales' indicators in the model specification.
*** $p < .001$.

both judged by human raters and processed using an AWE system (Burstein et al., 2020). The AWE system generated a variety of computationally-derived features that represent finer-grained aspects of writing construct subcomponents (Burstein, McCaffrey, Klebanov, & Ling, 2017). Thirty-six finer-grained features are used in the study. Fig. 2 illustrates the subconstructs associated with these features as well as a subset of some of the subconstructs. The set of 36 features were standardized and combined using a principle component approach into six components: argument (2 sub-features), organization and development (7 sub-features), reflection (3 sub-features), sentence structure (7 sub-features), vocabulary (10 sub-features), and conventions (7 sub-features).

7. Analyses

We first conducted a descriptive analysis and checked the mean, SD, and distribution of each survey question across study participants. We then conducted an exploratory factor analysis (EFA) within each of the four subscales based on a theoretical model specified in Philippakos and MacArthur (2015) and MacArthur et al. (2016): writing goals, writing efficacy, writing beliefs, and writing affect. The EFA results were used to determine an internal structure for each of the four subscales. In other words, a measurement model that linked the students' scores on the subset of survey items in this subscale to a latent factor was determined as a result of the EFA, after removing the factor loading values that were below .30 (Comrey & Lee, 1992; Costello & Osborne, 2005). The measurement model specified was then fit to the data for each subscale separately using confirmatory factor analysis (CFA). All analyses were completed using MPLUS 7.31 (Muthén & Muthén, 2015). Data model fit was evaluated using criteria suggested by Hu and Bentler (1999).

Subscale scores (observed scores based on items belonging each subscales as a result of CFA) were then used to calculate the Pearson product moment correlation with a series outcomes variables to explore the relationship between subconstructs of the writing motivation survey and AWE feature-based components, as well as college GPA, admissions test scores, and standardized writing assessment scores All correlational analyses were conducted using SPSS 23 (IBM Corp, 2015).

8. Results

Results are first presented in terms of each of the scales used in the survey. We then turn to the results of the survey as they are related to academic performance indicators.

8.1. Writing goals

The 11 items were intended to measure three goal orientations toward writing: mastery goals, performance goals, and avoidance goals. These items were taken from Philippakos and MacArthur (2015) after removing three items about getting good grades were deleted. Exploratory analysis indicated that the 11 items might be measuring the three intended types of goals: mastery, performance, and avoidance. Following suggestions from Comrey and Lee (1992) and Costello and Osborne (2005), we removed some items either with very low values of loadings or unspecified loadings. The items and loadings of the writing goals subscale are shown in Table 2.

Table 3
Writing Confidence Items.

	Mean	SD	Loading
1. I can write a paragraph with a clear topic sentence.	76.80	18.89	.73***
2. I can write complex sentences without making grammatical errors.	68.41	20.57	.71***
3. I can set goals for improving my writing.	74.47	19.89	.66***
4. I can think of a lot of ideas for my writing.	72.35	21.80	.70***
5. I can write a well-organized essay with an introduction, body, and conclusion.	76.17	19.61	.78***
6. I can evaluate whether I am making progress in learning to write.	71.06	20.22	.69***
7. I can write a paper using correct grammar.	75.37	19.24	.69***
8. I can organize paragraphs with ideas to support the topic sentence.	78.57	16.95	.79***
9. I can think of many words to describe my ideas.	71.02	19.63	.76***
10. I can plan before I write using an outline or organizer.	71.89	22.62	.58***
11. I can use punctuation correctly in all my sentences.	75.34	19.64	.63***
12. I can avoid distractions while I write.	58.16	25.57	.59***
13. I can end an essay with a strong conclusion.	67.12	22.36	.72***
14. I can come up with original ideas for my writing.	75.55	19.24	.76***
15. I can use commas and semi-colons correctly in my sentences.	68.94	21.31	.63***
16. I can plan time to get my writing done by the deadline.	77.10	21.87	.63***
17. I can start an essay with an interesting introduction.	72.10	20.34	.79***
18. I can focus on my writing for at least one hour.	74.58	22.33	.62***
19. I can find ideas to write about when I'm given a topic.	76.04	19.17	.78***
20. I can write a paper without spelling mistakes.	72.60	23.22	.62***
21. I can think of the perfect words to express my ideas.	65.85	20.27	.76***
22. I can tell when to use different writing strategies.	64.77	21.70	.70***
Cronbach's alpha est.		.95	

*** $p < .001$.

Table 4
Writing Beliefs Items and Subscales.

	Mean	SD	Content	Conventions
1. Writing helps make my ideas clearer.	3.81	0.90	.72***	
2. Revising is mostly about fixing errors in my grammar.	3.64	1.04		.28***
3. Writing helps me think about my topic in a new way.	3.78	0.85	.71***	
4. Good writers do not make errors in spelling.	2.45	1.19		.56***
5. The main problem of poor writers is using incorrect grammar.	2.80	1.08		.53***
6. I learn new things from writing.	3.84	0.91	.80***	
7. Good writers discover new ideas while writing.	4.15	0.73	.69***	
8. Writing quickly is an important part of good writing.	2.39	1.11		.62***
9. Good writers need little revision because they get it right the first time.	2.25	1.15		.75***
10. Good writers have to be able to write long sentences correctly.	3.34	1.01		.50
11. Writing is one of the best ways to explore new ideas.	3.85	0.90	.74***	
12. Revising helps me clarify my ideas.	4.18	0.78	.65***	
Correlation (Content, Conventions)			.19*	
Cronbach's alpha est.			.82	.66

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

Note three items have loadings on both performance and avoidance goals.[1]

CFA results suggested that this model (as shown in Table 2) fits the data acceptably well,: with $\chi^2 = 189.58$, $df = 38$, $p < .001$, RMSEA= .08, 90% CI = (.07, .10), CFI = =.94..96, TLI = Cronbach alpha coefficients were all above .72. Based on the model, the latent factors for the mastery and avoidance subscales were not significantly correlated. However, the factor for mastery correlated moderately with performance, and performance correlated with avoidance at a low level.

8.2. *Writing confidence (self-efficacy)*

For the writing confidence (self-efficacy) scale, the 1-dimension model—suggested in MacArthur et al. (2016) —fits the data marginally acceptably, with $\chi^2 = 862.87$, $df = 152$, $p < .001$, RMSEA = .091, 90 % CI = (.09, .10), CFI = .89, TLI = .88. The writing confidence items and their loadings are shown in Table 3.

[1] We also fitted the 3-factor model identified in MacArthur et al. (2016) where items 1 and 4 were loaded on both mastery and performance subscales, but this model had a poor fit to the data.

Table 5
Writing Affect Items.

	Mean	SD	Loading
1. I usually enjoy writing.	3.31	1.17	.92***
2 [a]. I don't like to write.	3.21	1.21	.90***
3. The process of writing is satisfying for me.	3.13	1.08	.81***
4. I think that writing is interesting.	3.54	1.07	.85***
5 [a]. I try to avoid writing as much as possible.	3.19	1.19	.74***
Cronbach's alpha est.		.90	

[a] Items 2 and 5 are reversely coded due to negative wording.
*** $p < .001$.

Table 6
Observed Intercorrelations Among the Seven Subscales Scores of the Writing Motivation Survey.

	Writing Goals: Mastery	Writing Goals: Performance	Writing Goals: Avoidance	Writing Confidence	Writing Beliefs: Content	Writing Beliefs: Convention
Writing Goals: Performance	.37**(.59)					
Writing Goals: Avoidance	.11*(.01)	.63**(.27)				
Confidence	.42**	.14**	−.26**			
Writing Beliefs: Content	.53**	.16**	−0.04	.47**		
Writing Beliefs: Convention	.06	.27**	.23**	.09*	.18**(.19)	
Writing Affect	.40**	.01	−.27**	.50**	.57**	.04

Note: numbers in parentheses are CFA estimated correlations among some of the subscales.
* $p < .05$.
** $p < .01$.

8.3. *Writing beliefs*

MacArthur et al. (2016) identified two sub-dimensions of the writing beliefs measure: belief in the importance of content and belief in the importance of conventions. The CFA analysis results suggested that the same 2-factor model defined in MacArthur et al. (2016) also fit the current data acceptably well, with $\chi^2 = 300.98$, $df = 53$, $p < .001$, RMSEA = .09, 90% CI = (08, .10), CFI = .922, TLI = .90. Cronbach's alpha coefficients were .82 and .66 for beliefs in content and in conventions, respectively. The two subscales correlated at a low level. The writing belief items and their loadings are shown in Table 4.

8.4. *Writing feeling (affect)*

The 1-dimension model fit the data adequately based on $\chi^2 = 233.76$, $df = 5$, $p < .001$, RMSEA= .29, 90% CI = (.26, .33), CFI = .97. The writing affect items and their loadings are shown in Table 5.[2]

8.5. *Relationships among the seven subscales*

Table 6 displays the observed intercorrelations among the seven subscale scores of the writing motivation survey based on the average scores of items in the subscales. The mastery goals subscale correlated with writing confidence ($r = .42$) and with writing content beliefs ($r = .53$), as well as writing affect (.40). Writing confidence correlated moderately with affect ($r = .50$) and beliefs about the importance of content ($r = .47$). Content beliefs correlated with affect ($r = .57$). Thus, correlations above .40 were found among five motivational subscales: mastery goals, confidence, content beliefs, and affect. Two negative correlations were found: Avoidance goals was negatively correlated with confidence (r =-.26) and with affect ($r = -.27$).[3]

[2] The EFA analysis also suggested that there might be two factors, one with items 1, 2, and 5, and another with items 1, 3, and 4. This model fit the data well, $\chi^2 = 3.27$, $df = 3$, $p = .35$, RMSEA = .01, 90% CI = (.00, .07), CFI = 1, TLO = 1. Items 2 and 5 present negative feelings toward writing while the others present positive feeling. This could contribute to the improvement in fit with a second factor. Also, the two factors correlated highly ($r = .75$). Consequently, we decided to combine these two factors in one.

[3] Note that the correlation of .63 between the performance goals scale scores and the avoidance goals scale scores was larger than the model estimate of the correlation between the latent performance goals and avoidance goals factors of .27 because, as shown in Table 2, three common items contribute to both the performance goals and the avoidance goals scale scores.

Table 7
Survey Results and Academic Performance Indicators.

Academic Performance Indicators	n	Goal: Mastery	Goal: Performance	Goal: Avoidance	Confidence	Belief: Content	Belief: Convention	Affect
Admission Scores: SAT								
SAT Reading and Writing Test (includes SAT or converted ACT)	274	−.05	.03	−.12*	.11	−.05	−.21**	−.02
SAT Math Test (includes SAT or converted ACT)	471	−.10*	.03	−.06	.03	−.15**	−.12**	−.15**
SAT Total Score (includes SAT or converted ACT)	438	−.07	.03	−.11*	.11*	−.08	−.19**	−.03
Admissions Scores: ACT								
ACT English Test	222	.03	−.01	−.21**	.25**	.06	−.22**	.09
ACT Writing Test	30	−.05	−.02	−.12	−.02	.06	−.27	.04
ACT Math Test	223	−.10	.03	−.08	.11	−.09	−.13	−.04
ACT STEM Test	223	−.10	−.00	−.14*	.12	−.10	−.17*	−.07
ACT Composite Score	225	−.02	.01	−.19**	.20**	−.01	−.21**	.06
Concurrent Academic Measures								
HEIghten™ Critical Thinking Test	324	−.04	.00	−.05	.06	−.04	−.19**	−.05
HEIghten™ Written Communication Test	325	.01	.03	−.06	.09	.00	−.15**	.02
HEIghten™ Written Communication Test – Non-Essay Score	325	−.02	.02	−.06	.10	−.05	−.17**	−.02
HEIghten™ Written Communication Test – Essay Score	325	.11	.06	−.04	.07	.12*	−.03	.15**
Academic Measures								
Study Semester Cumulative GPA	563	−.03	.08	.02	.04	−.06	−.06	−.02
Study Semester Plus One Semester Cumulative GPA	526	−.06	.06	.01	.07	−.06	−.03	−.02
Study Semester Plus Two Semesters Cumulative GPA	456	−.05	.07	-.01	.14**	−.02	−.01	.01
Study Semester Plus Three Semesters Cumulative GPA	364	−.04	.09	.05	.12*	−.01	.00	−.00
Study Semester Plus Four Semesters Cumulative GPA	215	−.00	.13	.11	.10	.07	.02	.01

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

8.6. Survey results and academic performance indicators

Table 7 provides correlations of the survey (average scores of items within each scale/subscale) with admission tests (SAT and ACT), concurrent outcomes assessment scores on the HEIghten™ Critical Thinking and Written Communicaton assessments, and GPAs of current and later semesters.

The column next to the label of outcome variables in Table 7 displays the sample sizes on which the correlation coefficients were computed. All of the statistically significant correlations are small (ranging from r =.00 to r = .25). Seven statistically significant positive correlations were found in the study, five were correlations of confidence with the SAT total score (r = .11), the ACT English (r = .25) and Composite (r= .20) scores; and GPA of the second (r = .14) and the third (r =.12) semesters following the survey administration. Two other positive correlations were between HEIghten Written Communication essay scores and writing beliefs in content (r = .12) and writing affect (r = .15). All other statistically significant correlations are negative. Beliefs in the importance of conventions correlated with over half of the performance measures including all SAT scores and 3 of 5 ACT scores. Avoidance goals were negatively correlated with five performance measures. Beliefs in conventions correlated negatively with written communication skills as measured by multiple-choice based items in the HEIghten Written Communication assessment (r = -.17).

8.7. Survey results and AWE features

Table 8 shows the correlations between motivation survey subscale scores and AWE feature components based on essays from the HEIghten Written Communication assessment.

Most of the significant correlations involve AWE feature components for conventions, which were positively correlated with

Table 8
Intercorrelations Between Writing Motivation Subconstructs and AWE Feature-Based Components.

AWE Feature Components	Goal: Mastery	Goal: Performance	Goal: Avoidance	Confidence	Belief: Content	Belief: Convention	Affect
Argument	−.00	.001	−.05	.03	.00	−.03	.02
Organization and Development	−.02	−.04	−.07	.00	.00	−.07	−.01
Reflection	.012	−.04	−.06	−.01	.12*	−.023	.08
Sentence Structure	−.04	−.01	.03	−.05	−.02	.07	−.04
Vocabulary	.00	−.00	−.12*	.15**	−.04	−.12*	.07
Convention	.15**	.06	−.14*	.18**	.15**	−.12*	.15**

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

mastery goals (r=.15), confidence (r =.18), belief in the importance of content (r=.15), and affect (r =.15), but negatively correlated with avoidance goals (r =-.14) and beliefs in the importance of conventions (r =-.12). The vocabulary feature component was negatively correlated with avoidance goals (r =-.12) and conventions beliefs (r =-.12) and positively with writing confidence (r =.15). The personal reflection feature component was only positively correlated with content beliefs (r =.12). Again all the correlations are small in absolute value.

9. Discussion

Our major goals of this research are focused on two research questions based on (1) the internal structure of a writing motivation survey and (2) the relationship between writing motivation subconstructs and a series of academic performance indicators, including features of student writing. Our discussion focuses on our findings as they are consistent with theory and prior research.

9.1. Internal structure of the motivation survey

Dimensional analyses of our more representative sample of college students showed the internal structure of the four survey scales as similar or identical to the structures reported by MacArthur et al. (2016). Writing goal orientation was fit to a 3-factor structure, and writing beliefs was fit to a 2-factor structure. The writing confidence (self-efficacy) and writing affect scales were each fit to a unidimensional structure.

Findings regarding goal orientation are consistent with theory (Senko et al., 2011) and prior research; two studies of achievement goal orientation in writing (Pajares & Cheong, 2003; Pajares et al., 2000) found the same three factors. For self-efficacy, some prior research has found separate factors for composing and skills (Pajares, 2007), and one study (Bruning et al., 2013) found three factors for ideas, conventions, and self-regulation. However, more research has used a single self-efficacy factor (Pajares & Valiante, 2006).

Furthermore, the correlations among the motivation constructs are consistent with theory. Mastery goals, confidence, affect, and beliefs in the importance of content were all moderately correlated ($r > .40$). A complementary pattern of negative correlations of avoidance with confidence and affect was also found. The same patterns of correlations were found by Philippakos and MacArthur (2015), with one exception (no correlation between mastery goals and confidence). Students who are confident in writing are more likely to focus on mastering writing tasks, find writing satisfying, and see writing as a meaningful task (content beliefs). Those who lack confidence in writing appear to be more likely to avoid writing and do not enjoy it or find it satisfying.

9.2. Relationship between writing motivation and academic performance indicators

The study also explored relationships between writing motivation and multiple indicators of outcomes and achievements. The study went beyond prior research by investigating longitudinal effects on GPA and fine-grained analysis of specific features of writing derived from AWE. The study found positive correlations between writing confidence and several achievement measures, including the SAT and ACT total scores, the ACT English test, and even with longitudinal outcomes of GPA two and three semesters later. Most research on writing self-efficacy has found positive correlations with writing measures (Pajares & Valiante, 2006; Philippakos & MacArthur, 2015). The current study shows correlations that were notably lower than those found in other studies. While correlations with measures of overall academic performance were low, they were significant. In the current study, the surveys were not necessarily conducted during students' regular class time. This timing might have impacted students' attention to the surveys and weakened the connection between the survey responses and their class writing activities, which might have resulted in less reliable motivation scale scores and weaker correlations. Nevertheless, the patterns in the correlations do align to previous research as further reviewed below. Even college admission tests such as the SAT, designed to predict college readiness and success, only correlate with first-year GPA between .20s and .30s (Beard & Marini, 2018), a level that is only slightly higher than those found in this study. For the significant but low values of correlations between writing motivation components and GPAs two or three semesters after the study semester, one could argue that these correlations may point to writing motivation specifically or motivation in general as an important factor of college success. Moreover, as found in previous research (e.g. Lei, Bassiri, & Scholtz, 2001), GPA (cumulative GPA by the end of each semester or first year GPA) is also not a perfect measure of college academic performance, even though multiple instructors or faculty members contributed to the GPA.

The study found consistent patterns of small negative correlations between standardized achievement tests and both avoidance goals and convention beliefs. The findings about avoidance goals are consistent with prior research on achievement in general academics (Senko et al., 2011) and in writing (MacArthur et al., 2016; Pajares et al., 2000); the pattern is the mirror image of the positive relationship between confidence and achievement. The findings about avoidance goals can resonate with prior research on writing apprehension, which is logically related to avoidance of writing and (negatively) to confidence (Daly & Miller, 1975) and writer's block (Rose, 1980).

The current findings of negative correlations between achievement and "conventions beliefs" are consistent with the two prior studies that investigated those beliefs (MacArthur et al., 2016; Philippakos & MacArthur, 2015). MacArthur et al. (2016) found significant negative correlations between "conventions beliefs" and all six writing outcome measures including essay quality (ranging from r = -.32 to r =-.47). Philippakos and MacArthur (2015) only included a single achievement measure, quality of a single essay, but they also found a negative correlation with "conventions beliefs" (r =-.14). The negative correlation with belief in conventions indicates that writers with lower performance measures have stronger beliefs that knowledge of conventions is a critical aspect of writing. This focus on the importance of conventions may be due to the higher salience of conventions for students who struggle the

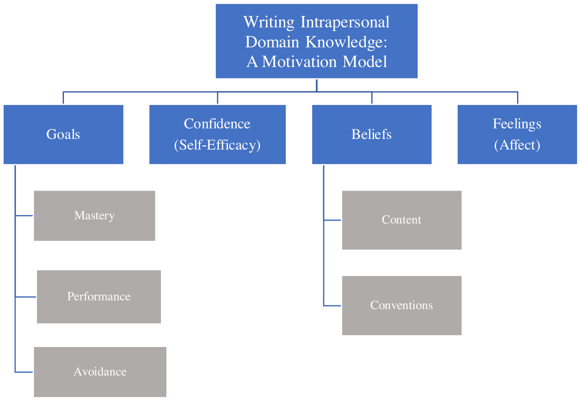

Fig. 1. Writing Intrapersonal Domain Knowledge: A Motivation Model.

most with basic writing.

Additional findings confirm the importance of learning more about writing conventions as perceived by students. In the present study, a statistically significant negative correlation (r =-.15) was identified in Table 7 between beliefs related to conventions and the HEIghten™ Written Communication assessment—the only concurrent academic measure that included a writing sample. It is true that only one of the four dimensions of the assessment—knowledge of language use and conventions—involves assessment of Standard Written English. Nevertheless, mean scores on the survey conventions items indicate that, overall, students disagreed with an emphasis on conventions, while weaker writers did not disagree enough. Table 4 also shows that students in the present sample disagreed with, or were neutral in, their reactions to such statements as "Good writers do not make errors in spelling" (Question 4) and "The main problem of poor writers is using incorrect grammar" (Question 5). The negative correlation may therefore be interpreted as an expected trend and, indeed, we might hope that the negative correlations would be higher. Pedagogically, we would want students to strongly disagree with each of the questions in Table 4 regarding knowledge of conventions. Students simply did not disagree strongly enough. As discussed below, writing instruction focused on the intrapersonal domain can allow students to realize that many of their commonly held beliefs about writing have been demonstrated as incorrect—especially in terms of the singular fixation on knowledge of conventions in teaching and assessing writing (Graham & Perin, 2007). The identification and use of the 2-factor structure related to writing beliefs thus lends a greater level of precision to interpretation of results.

Analysis of correlations between the modeled motivation constructs and specific features of writing identified by AWE are examined for the first time in this study. Statistically significant correlations were identified between the AWE features of reflection, vocabulary, conventions, and the motivation constructs identified in the survey. The findings for avoidance goals, conventions beliefs, and confidence are similar to the results with broader achievement measures. Avoidance goals and conventions beliefs are negatively correlated with vocabulary and conventions features. Confidence is positively correlated with vocabulary and conventions features. As with broader achievement measures, the correlations between motivation scale scores and features are small; however, their consistency with the results for the broader achievement measures and with the literature lends support for their interpretation as indicative of a meaningful relationship between motivation and writing performance.

10. Study limitations and future research directions

The study findings yield insights about the internal structure of a writing motivation survey and writing motivation's relationship with student success indicators and writing domain knowledge. We identify three limitations to our study and suggest directions for further study based on these limitations.

First, the number of institutions and students included in this study is very small in relation to the whole population of U.S. colleges and students. The inclusion of six institutions makes the findings related to both research questions more generalizable than previous research on this topic. Nevertheless, findings may be limited to those institutions and institutions similar to the six included in this study. Moving forward, expanded sampling plans would allow researchers to collect further sources of evidence. As well, larger sampling plans would support closer examination of subgroups based on race/ethnicity, gender assignment and gender identity, linguistic background, disability, age, socio-economic differences, and combinations of these and other categories. In turn, access to larger subgroup samples would allow researchers to determine if inferences are of comparable meaning across groups.

Second, this is the first study of its kind to examine relationships among motivation, writing features, and related performance indicators; therefore, our study is foundational and descriptive. Further experimental research will be needed that uses random assignment design to determine if factors other than designated interventions affect outcomes.

Third, we acknowledge that the instruments we have used to represent given constructs are limited. Both performance and motivation are multifaceted, as Wright et al. (2019) note in addressing the limits of their survey. Although the current study examined multiple theoretically important aspects of motivation, there were also aspects that were not addressed, including interest (Hidi &

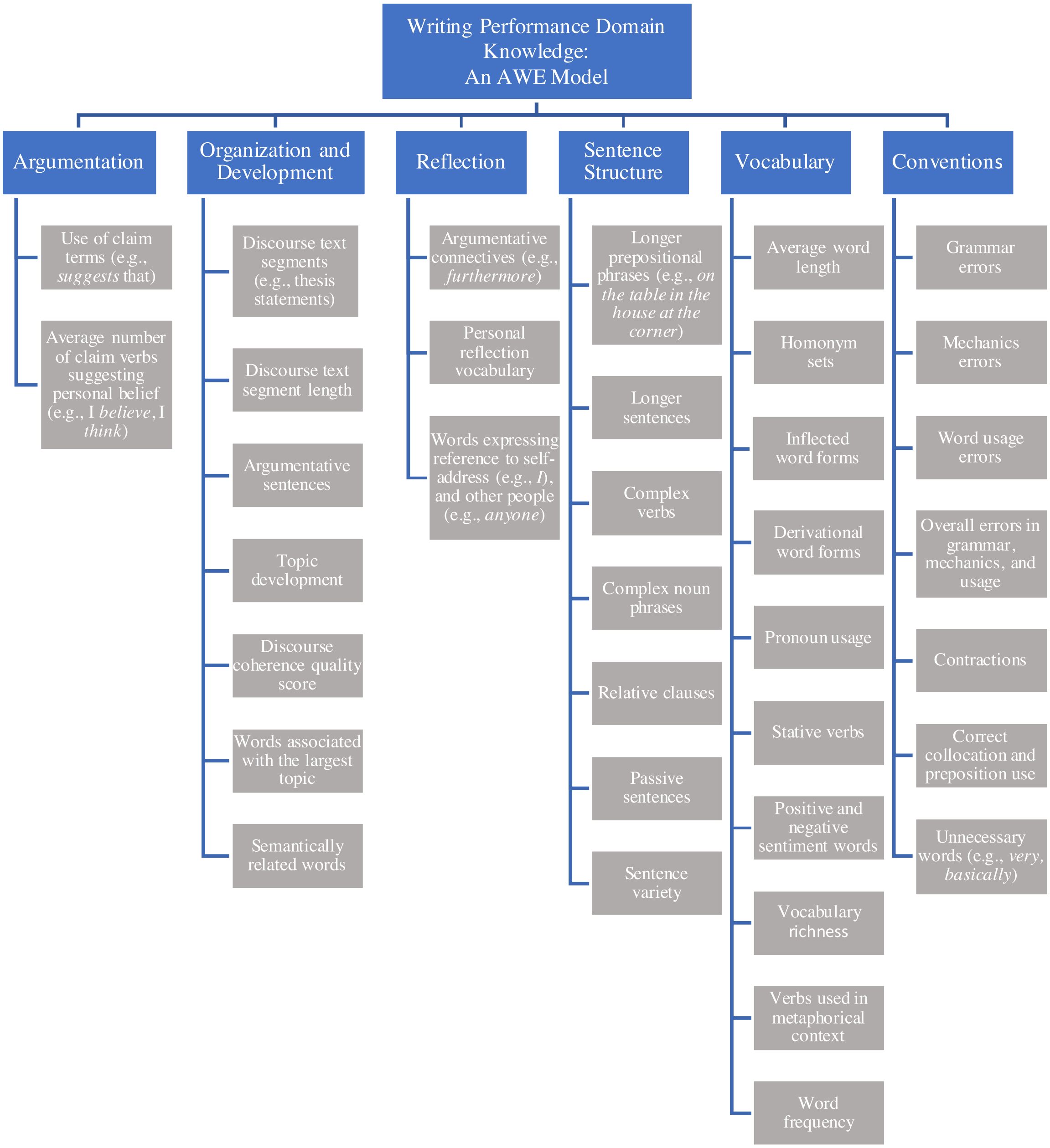

Fig. 2. Writing Performance Domain Knowledge: An AWE Model.

Boscolo, 2006), writing apprehension (Daly & Miller, 1975), and writer's block (Rose, 1980). Similarly, the AWE features used in this study were selected from several hundred features in the NLP feature portfolio available to the researchers (Burstein et al., 2020). Depending on genre, other features may be required to identify other relational evidence between motivation and the student success measures used in the present study.

11. Instruction in motivation and performance

Linking research to pedagogy, as Slomp (2020) has noted, is especially important in the present moment as we consider consequential perspectives of writing assessment that can be used to support instruction. As the United States continues to shift from routinized accountability to a culture of evidence, research on motivation will become increasing significant. Consistent with our findings, Bruning and Kauffman (2016), Pajares and Valiante (2006) and Wright et al. (2019) have found significant relationships between motivation and writing achievement. Our current study extends those findings to broader performance measures. Evidence suggests that students who are motivated are likely to produce better outcomes; therefore, instruction in writing should explicitly address writing motivation in addition to traditional focus on writing performance. In addition to discourse knowledge, skills, and strategies, students need to develop motivation to engage in writing in order to continue to improve and to use what they have learned in new situations.

One way to address both motivation and performance in instruction is to engage in construct-driven pedagogical practices. As Graham (2018) had suggested, conceptualization is key to pedagogy as it provides a visual road map of what is intended and what can be accomplished. Based on study findings, a model of the intrapersonal domain of writing is provided in Fig. 1. Based on the 36 AWE-based features used in the present study provided in Fig. 2, we also propose a model of subconstructs and associated features. Table 8 illustrates associations between writing motivation and the subconstructs of conventions, vocabulary, and reflection. Further research will be needed to provide increased understanding of the measurement properties of this feature-component relationship, as well as specific writing features within the three subconstructs and attitudes toward writing. Until that research is reported, both models may be considered for pedagogical use that combines explicit instruction in both motivation and writing feature use in terms of the overall constructs and their related subconstructs (Graham, Bruch et al., 2016; Graham, Harris, & Chambers, 2016).

A proven pedagogical approach that accomplishes these ends is strategy instruction with self-regulation (Harris & Graham, 2009). Systematic, explicit instruction in writing strategies provides the scaffolding needed by learners unsure about how to engage in planning, revising, critical reading, and other related language arts tasks. The effects of strategy instruction are enhanced by adding instruction in metacognitive strategies for self-regulation. As students engage in metacognitive self-regulation strategies—setting goals, managing tasks, monitoring use of strategies, and reflecting on what helped them learn—their self-efficacy increases. In turn, enhanced self-efficacy motivates further efforts to improve and to continue to use and improve on the knowledge and strategies that helped them succeed. A large body of research has demonstrated that strategy instruction in writing has strong effects on writing performance and that it has even stronger effects when it is integrated with instruction in self-regulation strategies (Graham, Harris et al., 2016; Graham & Perin, 2007).

At the college level, two studies (MacArthur, Philippakos, & Ianetta, 2015; MacArthur, Traga-Philippakos et al., 2019) have investigated strategy instruction with self-regulation. College basic writers learned genre-based writing strategies for rhetorical analysis, planning, and revising integrated with self-regulation strategies—goal setting, task management, progress monitoring, and reflection. To encourage beliefs that writing serves meaningful purposes, instruction focused on idea generation, organization, and clarity of expression with conventions considered in final editing. Furthermore, a strong emphasis was placed on learning self-evaluation so that students could evaluate their own work and see their progress. From a motivational perspective, the instructional goal was for students to believe that their success was due to using effective strategies, taking control of their learning, and effort. Self-evaluation and reflection on one's progress are critical to developing a growth mindset (Yeager & Dweck, 2012) that learning is possible with effort and strategic choices. Both studies found strong positive effects on the quality of argumentative essays as well as positive effects on motivation, including self-efficacy, beliefs, and affect.

Returning to Fig. 1, we can see how such explicit pedagogies can be formed in order to address the intrapersonal domain of writing as related to feature use as shown in Fig. 2. With the aim of advancing individual student opportunity to learn, those who teach and assess writing will find that research as that reported here—that is, research targeting performance and motivation—will provide an evidence base for new pedagogies.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

Research presented in this paper was funded by the Institute of Education Science, U.S. Department of Education, Award Number R305A160115. Any opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect the views of the IES. We are grateful to our partner institutions who made this work possible: Bloomsburg University of Pennsylvania, Bowie State University, California State University- Fresno, Jacksonville State University, Slippery Rock University, and University of North Carolina-Wilmington.

References

Abdel Latif, M. M. M. (2019). Unresolved issues in defining and assessing writing motivational constructs: A review of conceptualization and measurement perspectives. *Assessing Writing, 42*, Article 100417. https://doi.org/10.1016/j.asw.2019.100417

Allen, L., Dascalu, M., McNamara, D. S., Crossley, S., & Trausan-Matu, S. (2016). Modeling individual differences among writers using ReaderBench. In *EDULearn16: 8th international conference on education and new learning technologies* (pp. 5269–5279).

Allen, L. K., Snow, E. L., Crossley, S. A., Jackson, G. T., & McNamara, D. S. (2014). Reading components and their relation to the writing process. *L'année psychologique/Topics in Cognitive Psychology, 114*, 663–691. https://doi.org/10.4074/S0003503314004047

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

Bazerman, D. (2016). What do sociocultural studies of writing tell us about learning to write? In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23). Guilford.

Beard, J., & Marini, J. (2018). *Validity of the SAT® for predicting first-year grades: 2013 SAT validity sample*. College Board. http://files.eric.ed.gov/fulltext/ED582459.pdf.

Beigman Klebanov, B., Burstein, J., Harackiewicz, J., Priniski, S., & Mulholland, M. (2017). Reflective writing about the utility value of science as a tool for increasing STEM motivation and retention – an AI help scale it up? *International Journal of Artificial Intelligence in Education, 27*(4), 791–818. https://doi.org/10.1007/s40593-017-0141-4

Bruning, R., & Horn, C. (2000). Developing motivation to write. *Educational Psychologist, 35*(1), 25–37. https://doi.org/10.1207/S15326985EP3501_4

Bruning, R., & Kauffman, D. F. (2016). Self-efficacy beliefs and motivation in writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 160–173). Guilford.

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology, 105*, 25–38. https://doi.org/10.1037/a0029692

Burstein, J., McCaffrey, D., Klebanov, B. B., & Ling, G. (2017, September 8). Exploring relationships between writing & broader outcomes with automated writing evaluation. In *Proceedings of the 12th workshop on innovative use of NLP for building educational applications* (pp. 101–108).

Burstein, J., Riordan, B., & McCaffrey, D. (2020). Expanding automated writing evaluation. In D. Yan, A. Rupp, & P. W. Foltz (Eds.), *Handbook of automated scoring: Theory into practice* (pp. 329–346). CRC Press.

Cahill, A., & Evanini, K. (2020). Natural language processing for writing and speaking. In D. Yan, A. Rupp, & P. W. Foltz (Eds.), *Handbook of automated scoring: Theory into practice* (pp. 69–92). CRC Press.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10*(7). http://pareonline.net/getvn.asp?v=10&n=7.

Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal, 28*(2), 420–459. https://doi.org/10.11139/cj.28.2.420-459

Daly, J. A., & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. *Research in Teaching of English, 9*, 242–249. https://www.jstor.org/stable/40170632.

De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*(4), 687–698. https://doi.org/10.1037/0022-0663.94.4.687

East, M. (2019). Editorial: 25 years of assessing writing. *Assessing Writing, 42*, 1–3. https://doi.org/10.1016/j.asw.2019.100422

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218–232. https://doi.org/10.1037/0022-3514.72.1.218

Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. *College English, 41*(1), 19–37. https://www.jstor.org/stable/376357.

Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist, 54*, 258–279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and Writing, 26*(1), 1–15. https://doi.org/10.1007/s11145-012-9395-2

Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., et al. (2016). *Teaching secondary students to write effectively (NCEE 2017-4002)*. National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/wwc_secondary_writing_110116.pdf.

Graham, S., Harris, K., & Chambers, A. B. (2016). Evidence-based practice and writing instruction: A review of reviews. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 211–226). Guilford Press.

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology Monograph Series II, 6*, 113–135. https://doi.org/10.1177/2F2372732215624216

Hidi, S., & Boscolo, P. (2006). Motivation and writing. In: C. A. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 144–157). Guilford.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

IBM Corp. (2015). *IBM SPSS statistics for Windows, version 23.0*. IBM Corp.

Kauffman, D. F., Zhao, R., Dempsey, M. S., Zeleny, M. G., Wang, S., & Bruning, R. H. (2010, May). Achievement goals' impact on 11th grade students' affect, self-efficacy, and writing achievement. In *Paper presented at the annual conference of the American Educational Research Association*.

Lei, P., Bassiri, D., & Scholtz, E. M. (2001). Alternatives to the grade point average as a measure of academic achievement in college. *ACT research report series (2001-4)*. ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2001-4.pdf.

Liu, O. L., Mao, L., Frankel, L., & Xu, J. (2016). Assessing critical thinking in higher education: The HEIghten™ approach and preliminary validity evidence. *Assessment and Evaluation in Higher Education, 41*(5), 677–694. https://doi.org/10.1080/02602938.2016.1168358

MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 24–40). Guilford.

MacArthur, C. A., Philippakos, Z. A., & Graham, S. (2016). A multicomponent measure of writing motivation with basic college writers. *Learning Disability Quarterly, 39*(1), 31–43. https://doi.org/10.1177/0731948715583115

MacArthur, C. A., Philippakos, Z. A., & Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *Journal of Educational Psychology, 107*(3), 855–867. https://doi.org/10.1177/0091552113484580

MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? *Reading and Writing: An Interdisciplinary Journal, 32*, 1553–1574.

MacArthur, C. A., Traga-Philippakos, Z. A., May, H., & Compello, J. (2019, March). Self-regulated strategy instruction for basic college writers: Results from a randomized experiment. In *Paper presented at the annual conference of the society for research on educational effectiveness*.

Matsumura, L. C., Correnti, R., Zook-Howell, D., Walsh, M., & Bickel, D. D. (2020, January). Online content-focused coaching for high-quality reading instruction. In *Paper presented at the annual principal investigators meeting, institute of education sciences*.

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., et al. (2019). *The condition of education 2019 (NCES 2019-144)*. U.S. Department of Education. National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019144.

McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Mislevy, R. J. (2018). *Sociocognitive foundations of educational measurement*. Routledge.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (7th ed.). Authors https://www.statmodel.com/download/usersguide/MplusUserGuideVer_7.pdf.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578. https://doi.org/10.3102/00346543066004543

Pajares, F. (2007). Empirical properties of a scale to assess writing self-efficacy in school contexts. *Measurement and Evaluation in Counseling and Development, 39*, 239–249. https://doi.org/10.1080/07481756.2007.11909801

Pajares, F., & Cheong, Y. F. (2003). Achievement goal orientations in writing: A developmental perspective. *International Journal of Educational Research, 39*, 437–456. https://doi.org/10.1016/j.ijer.2004.06.008

Pajares, F., Britner, S. L., & Valiante, G. (2000). Relation between achievement goals and self-beliefs of middle school students in writing and science. *Contemporary Educational Psychology, 25*, 406–422. https://doi.org/10.1006/ceps.1999.1027

Pajares, F., and Valiante, G. (2006). Self-efficacy beliefs and motivation in writing development. In: C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 158–170). Guilford.

Philippakos, Z. A., & MacArthur, C. A. (2015). Writing motivation: Validation of a measure for college writers. June. In *Paper presented at the annual conference of the international association for the improvement of mother tongue education*.

Pintrich, P. R. (2000a). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press.
Pintrich, P. R. (2000b). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology, 25*(1), 92–104. https://doi.org/10.1006/ceps.1999.1017
Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology, 37*(1), 8–25. https://doi.org/10.1080/01443410.2015.11364071136407
Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Lawrence Erlbaum.
Rios, J. A., Sparks, J. R., Zhang, M., & Liu, O. L. (2017). Development and validation of the written communication Assessment of the HEIghten® Outcomes Assessment Suite. *ETS research report, RR-17-53* (pp. 1–16). https://doi.org/10.1002/ets2.12185
Rose, M. (1980). Rigid rules, inflexible plans, and the stifling of language: A cognitive analysis of writer's block. *College Composition and Communication, 31*, 389–401. https://www.jstor.org/stable/356589.
Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist, 46*(1), 26–47. https://doi.org/10.1080/00461520.2011.538646
Shaughnessy, M. P. (1977). *Errors and expectations: A guide for the teacher of basic writing*. Oxford University Press.
Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology, 87*(3), 386–398. https://doi.org/10.1037/0022-0663.87.3.386
Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology, 81*, 91–100. https://doi.org/10.1037/0022-0663.81.1.91
Shermis, M., & Burstein, J. (Eds.). (2013). *Handbook of automated essay evaluation: Current applications and new directions*. Routledge: Taylor & Francis Group.
Slomp, D. (2020). Volume 45 editorial. *Assessing writing* (p. 45). https://doi.org/10.1016/j.asw.2020.100472
Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research, 10*(2), 157–180. https://doi.org/10.1191/1362168806lr190oa
White, M. J., & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology, 30*, 166–189. https://doi.org/10.1016/j.cedpsych.2004.07.002
Wright, K. L., Hodges, T. S., & McTigue, E. M. (2019). A validation program for the Self-Beliefs, Writing-Beliefs, and Attitude Survey: A measure of adolescents' motivation toward writing. *Assessing Writing, 39*, 64–78. https://doi.org/10.1016/j.asw.2018.12.004
Yannakoudakis, H., Andersen, Ø. E., Geranpayeh, A., Briscoe, T., & Nicholls, D. (2018). Developing an automated writing placement system for ESL learners. *Applied Measurement in Education, 31*(3), 251–267. https://doi.org/10.1080/08957347.2018.1464447
Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 302–314. https://doi.org/10.1080/00461520.2012.722805
Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal, 31*, 845–862. https://doi.org/10.3102/2F00028312031004845

Guangming Ling is Managing Senior Research Scientist, Center for Education and Career Development, at ETS.

Norbert Elliot is professor emeritus of English at New Jersey Institute of Technology.

Jill C. Burstein is Director, Personalized Learning and Assessment Lab, at ETS.

Daniel F. McCaffrey is General Manager, Psychometric Analysis and Research, at ETS.

Charles A. MacArthur is professor in the School of Education at the University of Delaware.

Steven Holtzman is Principal Research Data Analyst in the Data Analysis and Research Technologies group at ETS.