Securing the right skills: A longitudinal assessment of college students' writing and public speaking self-efficacy

T. Kody Frey, University of Kentucky
Jessalyn I. Vallade, University of Kentucky

Abstract

This research investigated the developmental patterns of students' writing and public-speaking self-efficacy throughout their experience in the basic communication course (BCC). Questions were posed regarding (a) whether students grew in their reported writing and public speaking self-efficacy over two semesters, (b) whether growth differed based on biological sex, and (c) whether affinity and apprehension (as sources of performance self-efficacy) played a role in student growth. Two multilevel models revealed significant differences in students' initial status and rate of growth for each outcome. Specifically, sex, affinity, and apprehension influenced students' starting positions in the course, while only apprehension had a significant impact on growth. Analyses also revealed a contextual effect of the winter break on growth patterns. The results paint an important picture of the factors influencing students' personal development in the BCC, while also highlighting the use of multilevel modeling as a potential and relevant tool for contextualizing communicative development over time.

Keywords: self-efficacy, growth, basic communication course, multilevel modeling

As diversity steadily increases on college campuses (National Center for Education Statistics, 2019) and international crises highlight (and exacerbate) existing inequities in higher education (i.e., COVID-19), scholars face important calls to

reflect and grow, particularly regarding how they consider and address differential student experiences both within their classrooms and beyond (e.g., Ashby-King, 2021; Hampsten, 2021). Arguably one of the most far-reaching influences educators have in higher education is through the basic communication course (BCC). The BCC is often positioned as a general education requirement at many institutions, allowing it to transcend majors and touch the academic lives of students across campus (Valenzano et al., 2014).

Part of the reason BCCs have sustained their significance in higher education over time is a strong tradition of assessment. The National Communication Association (n.d.) recognizes the importance of this practice in its overview of the BCC within general education: "Assessment is valuable…because it identifies strengths, weaknesses, and areas for improvement. Although assessment can be a daunting task, it provides evidence that is useful when advocating for the resources that are needed to sustain a high-quality course" (para. 20). Essentially, communication scholars have effectively adapted multiple methods of assessment to ensure that the course is providing students with the skills they need for future employment opportunities. The National Association of Colleges and Employers (NACE) surveys employers annually who routinely rank written and oral communication skills as key characteristics desired in college graduates and employment applicants (Bauer-Wolf, 2019). For many students, the primary emphasis on and direct instruction for communication skill development comes from the BCC. Collectively, communication educators must ensure that they are facilitating skill growth equitably across a diverse student population, both for the immediate effectiveness of a given course experience and the potential long-term impact on career readiness and employment.

In pursuit of this idea, we see the need for new programs and methods of assessment reflecting increased diversity in terms of the courses being offered and the students enrolling in them. The present study builds on and contributes to these conversations by investigating students' performance self-efficacy in a two-semester, two-course BCC sequence. Performance self-efficacy has shown to be one of the strongest predictors of student achievement in higher education (Richardson et al., 2012; Schneider & Preckel, 2017). As a result, administrators should consider using the construct to operationalize behavioral learning in skill-focused courses and make arguments about effectiveness. That is, performance self-efficacy serves as an accessible measure of individuals' unique, personalized experiences that program administrators and scholars frequently use to argue that students are learning and can

perform the desired communication skills as a result of their experience in the BCC (e.g., Frisby et al., 2020; Strawser et al., 2017).

Self-Efficacy in the Introductory Communication Course

Rooted in social cognitive theory (Bandura, 1997), *performance* self-efficacy represents one's personal belief in their capability to accomplish academic tasks. Usher (2015) suggested conceptualizing performance self-efficacy in terms of *can-do* statements related to specific behaviors[1]. This is distinct from *academic* self-efficacy, which Richardson et al. (2012) defined as "general perceptions of academic ability" (p. 356). Whereas performance self-efficacy represents a belief based on students' experiences with domain-specific academic performances (e.g., "I can write a term paper of 15 to 20 pages," Shell et al., 1989, p. 99), academic self-efficacy involves more generalized representations of students' competencies (e.g., "How well can you learn reading and writing language skills?" Zimmerman et al., 1992). Both constructs are important predictors of academic achievement (e.g., Schneider & Preckel, 2017); however, given our interest in using self-efficacy as an indicator of student skill development, we are chiefly focused on evaluating performance self-efficacy.

Performance self-efficacy can be a valuable proxy for students' actual learning within the communication classroom (Strawser et al., 2017), and scholars have utilized this lens to forward important knowledge claims about effective communication instruction and the development of necessary communication skills (Dwyer & Fus, 2002; Nordin & Broeckelman-Post, 2020). However, as important as the information gleaned from these studies has been, knowledge claims about performance self-efficacy as a representation of student skill development within the BCC can still be extended.

Generally, within the field of communication, studies focusing on differences in performance self-efficacy based on identity factors such as sex have been inconsistent in their findings. Further, recent calls for scholarship that investigates the inclusiveness of classrooms and equity of learning outcomes (Faulkner et al., 2021; Nordin & Broeckelman-Post, 2020) supports the need to examine how students are growing in their general, foundational courses. As Nordin and Broeckelman-Post (2020) note, "if the introductory course truly serves as the front

[1] Unless otherwise noted, references to specific forms of self-efficacy in this manuscript (e.g., writing self-efficacy; public speaking self-efficacy) represent conceptualizations of performance self-efficacy. That is, we are interested in students' capabilities to perform domain-specific tasks rather than their general academic competence in the content area.

door, we must ensure that all students feel equally welcome to enter" (p. 99). Accordingly, the present study builds on these conversations to expand our understanding of students' experiences in the BCC. Following the recommendation of Broeckelman-Post et al. (2020), we utilize multilevel modeling (MLM; Hox, 2002) to determine (1) whether students are experiencing growth in their reported performance self-efficacy (writing and public speaking), (2) whether sex differences exist in student skill development, and (3) whether affinity and apprehension toward writing and public speaking (as psychological and emotional sources of performance self-efficacy) also play a role in this process.

Changes in Performance Self-Efficacy Over Time

First, it is important to note that not all BCCs are the same. Although some BCCs focus primarily on students' public speaking skills, others take a broader approach with their position in general education to enhance composition skills as well. Given this multimodal nature (i.e., academic domains related to both writing and speaking), students may be required to enroll in multiple classes across subsequent semesters. Put simply, students in different variations of the BCC may need more time to effectively develop oral communication skills alongside writing skills. This also means that students must navigate separate academic domains across a longer time period, which could potentially influence the extent to which the course fosters skill development. Though we know that students *can* develop writing and public speaking self-efficacy simultaneously throughout their BCC experience (Frey & Vallade, 2018), research can paint a more accurate picture of the BCC experience by more precisely and accurately modeling changes in performance self-efficacy over time while considering group and personal differences.

Indeed, several assessment studies use cross-sectional data to analyze performance self-efficacy at specific points in time. Assessments of classroom outcomes specifically within communication rarely model behavioral growth as a result of experiences, instruction, or study (Lane, 2017). There is a clear and pressing need to incorporate statistical tools within programmatic assessment that both enhance precision and correctly conceptualize outcomes as a change in behavior over time. If BCC course directors and administrators are using indicators of student development (like performance self-efficacy) to advocate for the effectiveness of their courses, then it makes sense that they would want to appropriately characterize this growth. Therefore, we employ MLM to ascertain the nuanced influences on

student growth over the course of an academic year, posing the following research question:

RQ1: Are students growing in their reported (a) writing self-efficacy and (b) public speaking self-efficacy in the BCC?

Sex Differences in Performance Self-Efficacy

Scholars have also long been interested in the role that group and context differences play in the development of student skills (Usher, 2015). Specifically, questions remain whether women and men differ in their reported performance self-efficacy as a result of their involvement in a variety of courses (Byars-Winston et al., 2017). This is largely because much of the research concerning sex and self-efficacy has investigated general perceptions of ability in an academic domain rather than the performance of specific behaviors within that context. For example, Huang (2012) found that across 187 different studies of academic self-efficacy, male students generally reported slightly higher self-efficacy than female students. Moreover, these differences were found to be context-specific, with female students scoring higher in language arts and music and male students scoring higher in mathematics and social sciences. Yet the differences were also small, prompting Huang to call for more longitudinal investigations of student experiences.

Other studies that do investigate task performance tend to echo this small effect or indicate little sex difference. Byrne et al. (2014) found that women and men somewhat differed in their accounting self-efficacy related to certain tasks, while Mamaril (2014) found no significant sex differences between female and male students' engineering self-efficacy. Research on performance self-efficacy in *communication* as a content domain has also failed to reveal significant differences in growth between men and women (Broeckelman-Post et al., 2020; Nordin & Broeckelman-Post, 2019). Thus, one might expect this lack of difference to persist when examining the development of specific student skills. However, there is a lack of evidence concerning performance self-efficacy in multimodal BCCs, where students may be tasked with meeting learning objectives in a variety of domains simultaneously (e.g., writing, public speaking).

A substantial amount of research concerning the relationships between writing and performance self-efficacy as a function of sex has been conducted in early educational contexts, including elementary, middle, and high school (Pajares et al., 2007). Research generally supports the position that women report more confidence in their writing than their male counterparts; "It is evident that, regardless of the

ratings that boys and girls provide on writing self-efficacy measures, girls consider themselves better writers than the boys" (Pajares, 2003, p. 149). The extent that male and female students differ in reports of public speaking self-efficacy is much less clear, but perhaps it can be inferred through speculation about students' experience with writing and public speaking as gendered academic domains (Cleary, 1996). According to Pajares (2003), differences in self-efficacy potentially occur due to students' views of academic domains as largely masculine or feminine. In more feminine domains, like writing, women are likely to see themselves as much more capable and confident in their beliefs.

Comparatively, there is currently no existing evidence to suggest the extent to which students view public speaking as a form of masculine or feminine discourse. In fact, Frey and Vallade (2018), in their study of public speaking self-efficacy in the BCC, found that 64.6% of students had no prior experience taking a class geared towards the development of public speaking skills. This may be contrasted with writing, which is likely a key component of all primary school curricula. Thus, despite conjecture that public speaking is a more masculine domain by some academics (Campbell, 1989), students in the BCC may not have developed gendered expectations for public speaking as an academic domain; they do not know whether women or men generally expect to succeed in this area.

Thus, self-efficacy research within the BCC finds itself in an interesting position. On one hand, some research from social science disciplines suggests that women may have more confidence in their performance self-efficacy beliefs than men, but these findings largely stem from results with younger student samples as opposed to college students. On the other hand, there is not enough empirical or theoretical evidence to suggest that men and women will differ at all in their reports of communication-related self-efficacy. This leads to our second research question:

RQ2: Does sex play a role in students' growth in (a) writing and (b) public-speaking self-efficacy in the BCC?

Sources of Communicative Self-Efficacy

Last, scholars have noted that increased levels of anxiety and apprehension prohibit the development of public speaking and writing skills (e.g., Daly, 1978). As such, evaluations of self-efficacy should also consider how these factors play a role in its development (Burns et al., 2021). Performance self-efficacy develops from four different sources: mastery experiences, vicarious experiences, social persuasions, and physiological and affective reactions (Bandura, 1997). Although previous research

has provided some background and context related to mastery experiences in the BCC (see Frey & Vallade, 2018), the current study highlights the influence of students' reports of affinity and apprehension (as conceptualizations of physiological and affective reactions) on reports of writing and public speaking self-efficacy.

Usher and Pajares (2008) noted that "strong emotional reactions to school-related tasks can provide cues to expected success or failure" (p. 754). Emotional reactions play a critical role in the development of students' communicative capabilities, prompting researchers to investigate factors that lead to the persistent development and sustainment of this response (Lefebvre et al., 2018). Researchers have fortunately demonstrated that, despite widespread prevalence of communicative anxiety among student (Bodie, 2010), effective intervention through the BCC can help students overcome this emotion and develop increased efficacy beliefs as a result (Broeckelman-Post & Pyle, 2017; Hunter et al., 2014; Lefebvre et al., 2020). As sources of self-efficacy shift and change over time, self-efficacy does as well, with students learning and gaining confidence in their communication skills. Taken together, a wealth of scholarship has focused on explicating the role of anxiety in communication courses as well as its impact on individuals' communication skills. However, other affective responses have received less attention in this context.

Much of the research on BCCs focuses on student learning outcomes, generally in the form of skill acquisition or behavioral indicators (e.g., Wallace, 2014). Though scholars have acknowledged the importance of the overall emotional climate of a BCC classroom for students' connectedness and retention (e.g., McKenna-Buchanan et al., 2020), affective responses related to the areas of focus have not been widely considered. In other words, while anxiety and apprehension have often been examined in domain-specific ways (e.g., public speaking anxiety), positive affect has not. Thus, in the present study, affinity for public speaking and writing, respectively, are examined in conjunction with anxiety responses to investigate these sources of performance self-efficacy more fully. This leads to the third research question:

RQ3: Do affect and apprehension play a role in students' growth in (a) writing and (b) public-speaking self-efficacy in the BCC?

Methods

Procedures and Instrumentation

The data reported in this study are derived from an assessment of the BCC at a large, Southeastern university from Fall 2018 and Spring 2019. The BCC consists of two courses – offered in consecutive semesters – that students choose to complete in pursuit of their general education requirements. Students are not required to take both courses successively, and they can choose to exit the BCC sequence to fulfill their core requirements elsewhere. Course instructors included a mix of graduate teaching assistants, part-time lecturers, full-time lecturers, and tenure-track professors, with some instructors teaching multiple course sections. All procedures were approved by the institutional review board.

As part of the course assessment, students are asked to complete a pre-test and a post-test each semester to earn course credit. Students complete the pre-test during the first two weeks of each course and the post-test during the final two weeks of each semester. Thus, the questionnaires are administered following a similar timeline from year to year: Wave 1 (Pre-Test; August), Wave 2 (Post-Test; November/December), Wave 3 (Pre-Test; January), Wave 4 (Post-Test; April/May). Thus, while all students enrolled in the course each semester are included as part of the sampling frame, the current research relies on data from 825 participants who completed at least two of the four questionnaires administered in the given year. Students who did not consent for their responses to be included as part of the analyses were excluded (n = 79 in Fall 2018; n = 58 in Spring 2019).

Dependent Variables

The procedures defined above include the collection of student reports of their (a) writing self-efficacy (WSE) and (b) self-perceived public speaking competence (SPPSC) across four waves. Each measure focuses on concrete tasks within the respective academic domain to reflect the focus on performance self-efficacy in the study. Hayes and Coutts' (2020) OMEGA macro was used to calculate McDonald's omega (ω) for reliability.

Writing Self-Efficacy

Writing self-efficacy was assessed using 7 items related to students' perceived capability for performing certain writing skills (see Frey & Vallade, 2018; Housley

Gaffney & Kercsmar, 2016; Strawser et al., 2017). Responses were collected using a 7-point Likert scale from *Strongly Disagree* (1) to *Strongly Agree* (7). Items included statements such as "I can properly cite sources in my writing" and "I can proofread my own writing for errors." The measure demonstrated reliability at Wave 1 (ω = .820; M = 4.97, SD = .90), Wave 2 (ω = .876; M = 5.40, SD = .91), Wave 3 (ω = .876; M = 5.20, SD = .90), and Wave 4 (ω = .896; M = 5.43, SD = ..93).

Self-Perceived Public Speaking Competence

Public speaking self-efficacy was operationalized using 4 items from Ellis' (1995) Self-Perceived Public Speaking Competence (SPPSC) scale. The items were selected based on a previous factor analysis of this scale using a similar population of students (Stewart et al., 2017; Stewart et al., 2019). The 4 items assess various dimensions of public speaking performance (e.g., "I have difficulty using appropriate gestures," "I use appropriate facial expressions," "I use language that is extremely clear," and "I have trouble articulating my words clearly"). Responses were collected using a 7-point Likert scale from *Strongly Disagree* (1) to *Strongly Agree* (7). Reliabilities for the measure were low across all four waves: Wave 1 (ω = .604; M = 4.68, SD = .92), Wave 2 (ω = .580; M = 4.96, SD = .87), Wave 3 (ω = .629; M = 4.82, SD = .88), and Wave 4 (ω = .646; M = 4.98, SD = .94).

Independent Variables

To investigate the relationship between student growth and sex, affinity, and apprehension, we included several student-level variables.

Student Sex

One item was used to assess students' biological sex. The final sample included 320 men (38.8%) and 505 women (61.2%). Responses were then dummy-coded so that men = 0 and women = 1.

Writing / Public Speaking Affect

Writing and public speaking affect were operationalized using an adapted version of Mottet and Richmond's (1998) Affective Learning Measure. The measure consists of four items measured on a semantic differential scale ranging from *Strongly Disagree* (1) to *Strongly Agree* (7), with one of the original items (Fair/Unfair) replaced (Useful/Useless) to better fit the context. The writing affinity measure demonstrated

reliability at Wave 1 (ω = .829; *M* = 6.00, *SD* = .1.07), Wave 2 (ω = .818; *M* = 6.13, *SD* = 1.05), Wave 3 (ω = .822; *M* = 6.11, *SD* = 1.04), and Wave 4 (ω = .827; *M* = 6.25, *SD* = 1.01). Likewise, the public speaking affinity measure demonstrated reliability at Wave 1 (ω = .808; *M* = 5.94, *SD* = 1.00), Wave 2 (ω = .790; *M* = 6.04, *SD* = 1.03), Wave 3 (ω = .814; *M* = 6.05, *SD* = 1.04), and Wave 4 (ω = .823; *M* = 6.26, *SD* = .95).

Writing Apprehension

Writing apprehension was measured using Autman and Kelly's (2017) six-item Writing Apprehension measure. Responses were collected using a 7-point Likert scale from *Strongly Disagree* (1) to *Strongly Agree* (7). Items included statements such as "I fear my writing being evaluated" and "Writing is not fun." The measure demonstrated reliability at Wave 1 (ω = .860; *M* = 3.71, *SD* = 1.29), Wave 2 (ω = .884; *M* = 3.47, *SD* = 1.41), Wave 3 (ω = .881; *M* = 3.60, *SD* = 1.33), and Wave 4 (ω = .901; *M* = 3.48, *SD* = 1.39).

Public Speaking Apprehension

Public speaking apprehension was assessed using a single dimension of the personal report of communication apprehension (PRCA-24; McCroskey, 1982). The dimension consists of six items addressing apprehension specifically within a public speaking situation (e.g., "Certain parts of my body feel very tense and rigid while giving a speech," "I face the prospect of giving a speech with confidence"). Responses were collected using a 5-point Likert scale from *Strongly Disagree* (1) to *Strongly Agree* (5). The measure demonstrated reliability at Wave 1 (ω = .813; *M* = 3.34, *SD* = .82), Wave 2 (ω = .822; *M* = 3.06, *SD* = .83), Wave 3 (ω = .824; *M* = 3.13, *SD* = .80), and Wave 4 (ω = .813; *M* = 2.91, *SD* = .77).

Given the longitudinal nature of the assessment procedures, multiple iterations of the dependent variables were nested within each student. This creates a hierarchical design that requires the construction of a series of multilevel models of individual change to control for violations of interdependence (Raudenbush & Bryk, 2002). Thus, a two-level MLM (i.e., four time points nested within students; Wave 1 [*N* = 559], Wave 2 [*N* = 502], Wave 3 [*N* = 672], Wave 4 [*N* = 672]) was used to analyze the data. We conducted separate analyses for each outcome (WSE and SPPSC) using Hierarchical Linear Models software (HLM8; Raudenbush et al., 2019).

Missing Data

When conducting a multilevel analysis, it is important to state whether data were complete and identify the methods used to account for any missing data (Dedrick et al. 2009). First, a notable number of students provided incomplete responses at Wave 2 (n = 80). Little's MCAR test was significant (χ^2 = 1548.352, p < .000), suggesting that the data may be missing at random (MAR) or missing not at random (MNAR; van Buuren, 2011). To further evaluate whether the missing data at Wave 2 was related to another variable in the dataset, we coded each observed variable as present or missing before conducting a series of independent samples t-tests. The t-tests revealed that missingness for several variables was related to other variables in the data. As an example, students who were missing data on the writing self-efficacy items generally had higher scores on the first item of the writing affinity measure (ts = -8.6 to -14.2, ps < .05).

An inspection of the incomplete responses revealed that all patterns of incomplete data could be attributed to six specific course sections (out of 29 sections offered total). Given our knowledge of the data and the assessment process used at this university, we are confident that missing data are due to the students' enrollment within these specific course sections. The instructors of these specific sections likely played a part in their students' incomplete responses, either by forgetting to remind them of the survey or by directly instructing them to complete some of the items rather than all of them so credit could be earned quickly. Moreover, 57 of the participants with missing responses at Wave 2 (68.8%) also had data included at Waves 1, 3, and 4. Thus, we used listwise deletion for all incomplete cases at Wave 2 but retained their responses across the other waves (van Buuren, 2011).

Next, we also evaluated the data for wave nonresponse (i.e., participants did not complete a wave or were late additions to the analyses) by conducting analyses of variance (ANOVA) with Bonferroni post hoc tests to examine mean differences across waves on the study variables. Across all variables, we identified four circumstances where students who responded to two waves differed from students who responded to four waves. The results imply that there are differences in wave completion rates among some key variables in the study and that the attrition in nonresponse across waves is due to the MAR mechanism. However, we believe MAR is a reasonable assumption for this longitudinal assessment data (Little, 2013). After one semester in the BCC sequence, many students choose to complete their remaining general education credits elsewhere. Or, some students choose to enroll

solely in the second half of this BCC sequence. Thus, some students in the sample are not presented with the opportunity to take part in the research at all four waves due to the nature of the course.

Following these procedures, no observed variables at any wave of data collection included more than 5% missing data (Tabachnick & Fidell, 2013). To preserve the sample size at each wave, remaining missing values were imputed using regression imputation with expectation-maximization. Finally, any students whose responses were only included at one wave of data collection were deleted.

Data Analysis

Statistically, we began the multilevel analysis by regressing the dependent variables on each time point for each individual, creating separate linear regressions to serve as baseline models (i.e., unconditional models)[2]: Importantly, no level-two predictors are included at this stage. For any parameters found to vary significantly (i.e., effects assumed to be different across students), we then expanded the model to include level-two predictors (e.g., sex, affinity, apprehension) of students' initial status (i.e., self-efficacy scores at Wave 1) and expected rate of growth (i.e., change in self-efficacy score across Waves 1-4). Student sex was left uncentered, the other independent variables were grand-mean centered (Kreft et al., 1995), and second-level effects were treated as fixed (i.e., assuming the effect of each predictor applies equally across students) in the analyses. Only the domain-specific apprehension and affinity variables were used to model the respective outcomes. Second-level variables were removed from the full models through backward stepping procedures; parameters with the largest *p* values were iteratively removed (Nezlek, 2008). The reported coefficients represented unstandardized beta weights (B), as is convention in multilevel modeling analyses and interpretation (see Hayes, 2006).

---

[2] For additional clarity, equations, and explanations of variables for all MLM analyses are available via the online supplement: https://osf.io/fjpsd/?view_only=3cf80645980a455d853b4c69f1f87d4a

Results

Prior to analyzing the research questions, we constructed a series of structural equation models to ensure that both measures of self-efficacy were factorially invariant over time (e.g., Putnick & Bornstein, 2016; Widaman et al., 2010). The goal in assessing longitudinal measurement invariance (LMI) is to ensure that an outcome represents "the same construct on the same numerical scale at each time point, and that difference scores reflect real gains or losses of [the outcome] over time" (Goodboy et al., 2021, p. 5). To account for multivariate nonnormality, models were constructed using longitudinal confirmatory factor analysis (CFA) with robust maximum likelihood estimation via the lavaan package (Rosseel, 2012) in the free statistical software R 4.1.0 (R Core Team, 2021)[3].

Measurement Invariance

Initially, using data from only the students who responded across all four waves of data collection (N = 349), we assessed the fit of the measurement model for each wave independently. Then, following the guidance of Widaman et al. (2010), tests for measurement invariance were conducted through three additional steps reflecting a series of increasingly restrictive models. The *configural* invariance model, or the least restrictive model, is used to ensure baseline equivalence of constructs across waves. If this model holds, *metric* invariance (also labeled weak invariance) is tested by setting factor loadings to be equal over time. Once metric invariance is established, *scalar* invariance (also labeled strong invariance) is then tested by setting both the factor loadings and intercepts to be equal over time. Invariance standards are evaluated based on changes in model fit between the configural and metric models, and between the metric and scalar models. To make these comparisons, we used the scaled likelihood ratio (LRT) chi-square statistic (Satorra & Bentler, 2001; $\Delta\chi^2$), along with several alternative fit indices (AFIs), including the Comparative Fit Index (ΔCFI), Tucker-Lewis Index (ΔTLI), and Root Mean Square Error of Approximation (ΔRMSEA; Putnick & Bornstein, 2016). All model fit indices and change statistics are reported in Table 1.

-- Insert Table 1 here –

[3] The R code for assessing longitudinal measurement invariance is also available via the online supplement: https://osf.io/fjpsd/?view_only=3cf80645980a455d853b4c69f1f87d4a

For WSE, results showed support for partial scalar invariance. When equality constraints were imposed upon the factor loadings, the model showed a significant decline in fit from the LRT (p = .017). Modification indices (MIs) indicated that freely estimating the loading for item 1 at Wave 3 (MI = 10.049; "I can adapt my writing to any given audience") would improve the overall fit. The resulting model did not show a significant decline in LRT (p = .140) and produced sufficiently small changes in AFIs between models (Chen, 2007; Cheung & Rensvold, 2002; Meade et al., 2008): ΔCFI (-.001), ΔTLI (.000), and ΔRMSEA (.000). This freed loading was retained when assessing strong invariance.

When equality constraints were imposed upon the intercepts (apart from the intercept of the one noninvariant loading; Putnick & Bornstein, 2016), the model again showed a significant decline in fit (p = .005). MIs indicated that model fit could be improved by releasing the constraints on several intercepts. Following the advice of van de Schoot et al. (2012), we released these intercepts iteratively – beginning with the largest MI – and retested the models for partial scalar invariance. After releasing two constraints (i.e., item 2 at Wave 2, MI = 7.079; "I can properly cite sources in my writing"; item 7 at Wave 4, MI = 6.386; "I can write essays that emotionally impact readers"), the model did not show a significant decline in LRT (p = .147) and had negligible differences in CFI, TFI, and RMSEA (ΔCFI = -0.001, ΔTLI = 0.000, and ΔRMSEA = 0.000) in comparison to the previous model. Thus, we concluded that partial scalar invariance was established for WSE, and valid inferences regarding the latent means could be made over time (Byrne et al., 1989).

For SPPSC, partial scalar invariance was also supported. The scalar invariance model showed a significant decline in LRT (p = .000), along with excessive changes in AFIs. Therefore, we released the constraints on two intercepts (i.e., item 4 at Wave 3 (MI = 9.320; "I have trouble articulating my words clearly"; item 3 at Wave 2 (MI = 7.006; "I use language that is extremely clear"). This resulted in a new model which did not show a significant decline in fit from the LRT model (p = .096) and produced negligible differences in CFI, TFI, and RMSEA (ΔCFI = -0.004, ΔTLI = 0.002, and ΔRMSEA = -0.005) in comparison to the previous model. We concluded that we could again make valid inferences regarding the latent means.

Research Questions

RQ1 sought to address whether students grew in their WSE and SPPSC from the beginning of the BCC course sequence to the end. The means of the dependent variables across each wave provide initial evidence for consistent growth across both

content areas. However, a graphical inspection of the means for both outcomes (Figure 1) demonstrates the presence of an unexpected dip in scores between Wave 2 (End November/Beginning December) and Wave 3 (January). Thus, students' reports of self-efficacy may not grow in a straight, linear line across waves due to a slump in perceived capability over the winter break.

-- Insert Figure 1 here --

Therefore, to account for potential losses in self-efficacy between courses in the BCC sequence and to more accurately model students' growth trajectories, we constructed two baseline models to test RQ1: (1) an unconditional linear growth model and (2) a linear growth model with winter break included as a time-varying covariate (TVC), which can be used to "model discontinuous growth in a longitudinal dataset" (see McCoach & Kaniskan, 2010, p. 2).

The unconditional model showed evidence of significant growth for both outcomes (Table 2). At the beginning of the BCC sequence, the average WSE score was 5.00, and it increased linearly at a rate of 0.15 per wave. The significant variance component for students' initial status at level-two also suggests that students began the course with different levels of WSE scores. Additionally, the variance component for rate of growth was not significant, indicating a lack of variance in growth rates over time (i.e., students are assumed to grow at the same rate). For SPPSC, the average score was 4.69 at the beginning of the course sequence, and it increased at a rate of 0.10 per wave. The variance components for both initial status and growth were significant, indicating that students began the course with different levels of SPPSC and grew at different rates over time.

-- Insert Table 2 here --

Estimating the effect of the winter break required the construction of a dummy-coded variable related to time before and after the break. Thus, data collected during the Fall semester were coded as 0, and data collected during the Spring semester were coded as 1. Winter break was also treated as a fixed variable; it assumes that students decrease over the period at a similar rate and takes on the same value for every person in the sample.

Results for the TVC model (Table 3) were almost identical to the unconditional model, with a few additions. First, the average WSE score at Wave 1 was 4.98, and it increased at a rate of 0.32 per wave. The average effect of the winter break was also significant (B_{20} = -0.47, $p < .001$). This means that, whereas students were growing an average of 0.32 points in WSE per wave, they were actually *losing* about 0.15 points (.32-.47) over the winter break. Likewise, the average SPPSC score was 4.68 at

Wave 1, with scores increasing at a rate of 0.22 per wave. The average effect of the winter break was again significant (B_{20} = -0.34, p < .001). Thus, whereas students were growing an average of 0.22 points in SPPSC per wave, they were losing about 0.12 points (.22-.34) over the winter break. Further, the addition of the TVC resulted in significant variance components at level-two. When accounting for the TVC, students began the course with different levels of WSE and SPPSC and grew at different rates over time. Second, when comparing the TVC model to the unconditional model, the residual variance accounted for at level-one decreased by 5.2% for WSE and 3.0% for SPPSC.

Collectively, these baseline models answer RQ1. Results revealed significant growth in both WSE and SPPSC during students' time in the BCC, despite a loss in performance self-efficacy occurring over winter break.

-- Insert Table 3 here --

Initial Status and Rate of Growth

RQ2 and RQ3 examined differences in WSE and SPPSC growth as a result of (a) student sex and (b) affect/apprehension toward writing and public speaking. To answer these questions, second-level predictors were added to the TVC models to account for the significant variation in both initial status and rate of growth. To get a better sense of the data, this first involved the examination of correlations between initial status and rate of growth. As described by Wilkins and Ma (2002):

> A negative correlation indicated a fan close pattern of change; that is, students with higher initial status grew less than did students with lower initial status. A positive correlation indicated a fan open pattern of change, in which students with higher initial status grew faster than did students with lower initial status, thus, creating greater variation in achievement over time. (p. 293)

For WSE, the relationship between initial status and growth was nearly non-existent (.04). This suggests that students who began with higher WSE grew at about the same rate as students who began with lower WSE. However, for SPPSC, the negative correlation between initial status and rate of growth was moderately strong (-.30); students who began the course with higher levels of SPPSC tended to grow less over time than students who began with lower reported levels. Because the variance components in Table 3 for WSE and SPPSC initial status and growth were

significant, predictors (student sex, writing affect, writing apprehension) were included to model this variation at the second level and add further context to these relationships.

Predicting Initial Status and Growth in Writing Self-Efficacy

Table 4 presents the results for the full MLM models.

-- Insert Table 4 here –

In the full model, the values for the intercept (4.90) and time slope (0.32) represent the initial WSE score and rate of growth, respectively, for male students with average writing apprehension and writing affect. When controlling for other variables in the model, women scored 0.17 points higher in WSE initial status than men, students who reported greater affect for writing were 0.12 points higher in initial WSE than students who reported low affect for writing, and students who reported greater writing apprehension were 0.36 points lower in initial WSE than students who reported less writing apprehension. Writing apprehension was also the only significant predictor of student growth. Students with higher writing apprehension scores generally grew faster in WSE than students with lower writing apprehension scores. For two students with writing apprehension one unit apart, students with the higher score grow .06 points faster than students with the lower writing apprehension score. Compared to the TVC model, the residual variance accounted for in WSE initial status decreased by 59.9%, and the residual variance accounted for in WSE rate of growth decreased by 2.1%

Predicting Initial Status and Growth in Public Speaking Self-Efficacy

The results for SPPSC mirror those for WSE. The values for the intercept (4.56) and time slope (0.22) represent the initial SPPSC score and rate of growth, respectively, for male students with average public speaking apprehension and public speaking affect. When controlling for other variables in the model, women scored 0.23 points higher in SPPSC initial status than men, students who reported greater affect for public speaking were 0.14 points higher in initial SPPSC than students who reported low affect for public speaking, and students who reported greater public speaking apprehension were 0.45 points lower in initial SPPSC than students who reported less public speaking apprehension. Like the results for WSE, public speaking apprehension was also the only significant predictor of student growth. Students with higher public speaking apprehension scores generally grew faster in SPPSC than students with lower public speaking apprehension scores. For two

students with public speaking apprehension one unit apart, students with the higher score grow .06 points faster than students with the lower score. Compared to the TVC model, the residual variance accounted for in SPPSC initial status decreased by 37.6%, and the residual variance accounted for in SPPSC rate of growth decreased by 21.1%.

The combined results answer RQ2 and RQ3. For RQ2, sex did not play a role in students' WSE or SPPSC growth, but it did influence students' starting positions in the course for each outcome. For RQ3, affinity and apprehension influenced students' initial status for WSE and SPPSC, but only apprehension subsequently affected growth.

Discussion

This study sought to provide a rationale for using self-efficacy as an indicator of student skill development over time in the BCC. We specifically tried to determine (1) whether students experienced growth in their writing and public speaking capabilities and (2) the role that student sex, along with affect and apprehension, played in that growth. We offer several conclusions based on the results.

First, students who took part in the year-long, two-course BCC sequence grew in their reports of both WSE and SPPSC. Students clearly improved in their perceived writing and public-speaking capabilities over time, providing evidence that the course is having a positive effect on perceived student skill development. Table 3 indicates this growth by painting a picture of the *average* student experience in the BCC. The average WSE at the beginning of the course was 4.98. From this starting point, the linear rate of WSE growth was 0.32 across each wave of data collection (i.e., the beginning and end of each course sequence). Then, the addition of the TVC indicated that students lost 0.15 points (.32 - .47) of WSE when they left campus for winter break. Likewise, the average score for SPPSC among students at the beginning of the course was 4.68, and the linear rate of growth for SPPSC was 0.22. The addition of the TVC indicated that students lost 0.12 (.22-.34) points of SPPSC when they left campus for winter break. Thus, despite this decline between courses, the typical student still appeared to leave the course sequence with higher levels of both outcomes than they had when it began. In terms of programmatic assessment, these results could be used to demonstrate the value of the BCC sequence, as students end their first year with increased perceptions of their writing and public speaking capabilities.

However, the results become more complex when investigating the nature of this growth among students. To begin, the variance components in Table 3 indicate that students were significantly different in their initial status for both outcomes. Practically, this means that students began the course at different levels of self-efficacy; some students may be entering the course in disadvantaged positions that could mitigate or harm their future development. The variance components in Table 3 also indicate that the nature of growth is different across students (i.e., students grow at significantly different rates). Ideally, students who begin with lower reports of WSE and SPPSC will grow faster than students who begin in more privileged positions to ensure equity of outcomes upon course completion.

Next, the correlations between initial status and rate of growth provide insight into the pattern of growth. The small, positive correlation for WSE (.04) suggests that students were slowly becoming dissimilar over time. However, this small effect, coupled with the small amount of variance captured at the second level, indicates that students' experiences in terms of their WSE are likely very similar. This may seem like an important, positive finding for the status of the course, as there is no evidence that some students are learning or developing in their writing more than others. But, if students grow at the same rate, then those who begin the course lower in WSE are bound to finish the course that way too.

Regarding SPPSC, there is evidence that students are indeed growing at different rates. The moderately strong, negative correlation between initial status and rate of growth in SPPSC (-0.30) implies that students were becoming increasingly similar in terms of their SPPSC. That is, students who began the course higher in initial SPPSC grew more slowly than students who began lower in initial SPPSC. Such differences highlight how the BCC experience may be affording students who begin the course with lower public speaking self-efficacy opportunities to have equitable outcomes with their peers by the time the course ends. Students who begin the course with lower capability beliefs may be able to catch up to the level of their peers by the end of the course sequence. The level-two variables (student sex, affinity, and apprehension) were introduced into each model to explain which students started at different levels and why they potentially grew at different rates. Researchers and administrators can only target, adapt, and adjust their courses once they pinpoint the gaps where students are clearly differing in the course experiences.

The full models (i.e., Table 4) demonstrate clear differences in students' WSE and SPPSC scores at the beginning of the BCC sequence when controlling for other variables in the model. For both outcomes, female students reported greater perceived capabilities at the beginning of the course than male students. This result contradicts previous research that failed to find significant sex differences in communication self-efficacy in the specific domain of public speaking (Broeckelman-Post et al., 2020; Nordin & Broeckelman-Post, 2019). Likewise, students who entered the course with higher levels of affect began with greater capability beliefs in each area, as did students with lower levels of apprehension. Such differences may point to a lack of equity in writing and public speaking opportunities prior to students' experiences in the BCC.

Consequently, continued consideration and empirical investigation of students' individual identities and differences in self-efficacy growth and outcomes within the

BCC will provide data and perspective about potential structural inequities, curriculum and assessment gaps, and instructional practices that may differentially impact student learning. If students are truly beginning their courses at different levels, then research should account and control for these individual differences. One particularly important area for consideration is the impact of race/ethnicity on student growth[4]. Emerging research suggests that students of various races and ethnicities may experience different mindsets related to changes and growth in communication skills and course objectives (Morreale et al., 2021; Nordin & Broeckelman-Post, 2020), though results have been inconsistent (e.g., Byars-Winston et al., 2017). For example, Nordin and Broeckelman-Post (2020) investigated differences in communication growth mindset, academic self-efficacy, and intercultural communication competence, discovering that Black students scored highest in growth mindset and efficacy, in contrast to research suggesting that Black students viewed themselves as having lower communication competence than the norm (Chesebro et al., 1992). These results alone demonstrate the need for updated research on race/ethnicity in communication and assessment research.

Yet we also remain concerned that such a call may be overlooked by communication scholars; issues of diversity and difference have long been acknowledged yet unaddressed in mainstream communication classroom research (e.g., Simmons & Wahl, 2016). If we truly want to improve students' developmental experiences, then research and scholarship must seek to better integrate and reflect students' individual identities and existing skill levels to achieve equitable outcomes. It is our hope that the evidence provided herein will motivate scholars to consider how issues of individual difference and identity impact applied assessment practices. As Morreale et al. (2021) noted, assessment efforts should not be considered as a

[4] Race/ethnicity was initially considered as a variable of interest for the current manuscript; however, there was not enough diversity in the sample to accurately model the influence of racial/ethnic differences on WSE and SPPSC. Students identified as White/Caucasian (n = 673; 82.0%), Black/African American (n = 65; 7.9%), Asian (n = 35; 4.3%), Native American (n = 2; 0.2%), biracial (n = 28; 3.4%), and other (n = 18; 2.2%). Thus, the sample was overwhelmingly White, and any attempt to empirically model group differences would likely have been underpowered. As such, claims derived from the analyses may have hidden true racial/ethnic differences that existed within the data; we wanted to avoid misrepresenting the experiences of various racial/ethnic groups through our statistical conclusions. Moreover, the entire sampling frame of students enrolled in the BCC course was given the opportunity to participate in the assessment; the reported percentages for race/ethnicity are likely similar to the true percentages across the entire sample. Race/ethnicity must be considered in future research so BCC administrators and schools can ensure equitable experiences in our courses. At the same time, our program is likely not the only one where small percentages of minority students complicate the use of quantitative assessment procedures. This may require unique and creative research solutions – like the use of purposive sampling that can ensure equal-sized groups – to ensure representation of the minority students in our courses.

mere data collection and should instead be an "impetus for continuous improvement" (p. 154).

Third, writing apprehension and public-speaking apprehension were both found to have a significant effect on students' growth in WSE and SPPSC. Students with more apprehension grew at a faster rate than students with less apprehension. Specifically, at each wave, the students with more apprehension surpassed the students with less apprehension by 0.06 points. Based on this result, we can conclude that the experience in the course is allowing students who began with greater fear of writing and public speaking to become more like their confident peers over time.

Yet, given the theoretical link between emotional responses and self-efficacy, we might have anticipated a stronger effect for apprehension or significant results for affinity. This is likely attributed to the measurement of these variables as operationalizations of the physiological and affective sources of self-efficacy. We collected data on students' physiological and affective reactions at singular points in time and used it to predict growth. Instead, better conceptualization of the variables as *reactions* to learning stimuli may fall more in line with the theoretical positions proposed within social cognitive theory and reveal significant findings. The stagnant, cross-sectional collection of the affective variables at a single point in time may be more suited to predict individualized instances of self-efficacy instead. Considering that self-efficacy also develops from (1) mastery experiences, (2) role modeling, and (3) vocal persuasion, other variables could also be used in the future to explain additional variance in student self-efficacy growth in this setting (Usher & Pajares, 2008).

Finally, the results point to the need for scholars interested in communicative skill development to use MLM to model the effects of individual differences and changes, as well as context, on student learning or classroom experiences. Researchers commonly use student-level characteristics and traits (e.g., race, socioeconomic status) as important controls before investigating contextual effects on classroom processes, and the current results reinforce why this should be more routinely practiced within communication. Scholars can use MLM to control for individual differences when longitudinally assessing their courses to paint a more accurate picture of students' BCC experiences (see Ma et al., 2008). Further, using the technique may add statistical precision to our investigations of classroom processes by revealing important or hidden contextual or group differences that go undetected when nested data is not properly analyzed. In this way, the use of MLM will provide an additional means for answering the calls of communication scholars

to contextualize student learning (Lane, 2017) and integrate considerations of the micro- and macro-levels of influence (Hampsten, 2021). As Lane (2017) asserted, "we are painfully aware that context matters, but we have largely overlooked how these important details impact instructional dynamics" (p. 121).

As an example, the current research revealed the contextual effect of the winter break as a feature of the two-course BCC sequence. Students demonstrated growth over the course of the first semester, only for that growth to be partially stunted by the loss of instruction that occurs over the winter break. Moreover, the effect of this break was even more pronounced when controlling for the individual predictors (-0.50 for WSE and -0.36 for SPPSC in the full model). When students return home from campus for the break, they are removed from the evaluative climate that defines the instructional setting (Nyquist & Booth, 1977). This sustained assessment differentiates the classroom from other contexts and is a fundamental component of how students reflect on their mastery experiences. Students use the feedback they receive in the form of grades or comments from instructors to construct authentic impressions of their capabilities, which in turn directly influence self-efficacy (Frey & Vallade, 2018). Consequently, the winter break effect provides evidence that student growth in the two-sequence BCC may not be linear. It may be helpful for instructors or administrators of this type of BCC to hold a brief review of important concepts at the beginning of the second course in the sequence so students can quickly retrieve information they may have lost over the break. Similarly, low-stakes communication activities and minor assignments that operate as skill sharpeners may help students regain some of their confidence and abilities that could have been dulled over the break.

Limitations and Future Directions

Several limitations should be considered when addressing the findings of this research. First, the methodological choices made when using MLM to answer the research questions have important implications; ignoring or disregarding levels can have serious consequences on the validity of the analyses (Bliese & Hanges, 2004; Huang, 2016; Huang, 2018). For example, the decision to specify a two-level model as opposed to more complex hierarchical structures changes the multilevel framework. Although we were primarily interested in assessing change (level-one effect) and the subsequent impact of specific variables on this change (level-two effects), students can also be nested within various other groupings (e.g., time of day, academic majors, specific classrooms, course sections, etc.). A three-level model may

be especially necessary in cases where the full population of students within a program is not available for assessment purposes. If students are randomly selected for analyses from groups of schools or instructors, then a third level may be an important control for the possibility of cluster effects resulting from those sampling procedures.

Notably, the current analysis excludes groups of students nested within instructors. Since instructors are known to have significant influences on classroom processes in higher education (Schneider & Preckel, 2017), responses from students who take classes from similar instructors are likely going to share similarities and violate regression assumptions of interdependence (Frey & Lane, 2021; Dresel & Rindermann, 2011). However, we recognize that many of the students in the current sample switch instructors when transitioning between the first and second courses in the BCC sequence. Thus, despite the importance of this structure, a more appropriate approach would involve third-level groupings that remain stable across the entire academic year (e.g., year in school). We urge scholars interested in MLM to strongly consider the implications of the groupings that they select or consult more complex forms of multilevel modeling (e.g., piecewise growth modeling, multilevel structural equation models) that are more flexible when dealing with complex data structures.

Next, despite the clear presence of growth because of students' experiences in the course, the results invite criticism over the usefulness of the two-semester, two-sequence model. Given the similarity in the mean self-efficacy scores at the end of Wave 2 (Fall semester) and the end of Wave 4 (Spring semester), it seems reasonable that BCC administrators might question the usefulness of the second-course sequence altogether. That is, there may not be any social or practical significance in having students enroll in two consecutive courses if they finish with the same beliefs in their abilities as students enrolled in a one-semester sequence. Since the loss of self-efficacy over the winter break was relatively large, we believe that the addition of the second course helps to inoculate students against potential larger losses in self-efficacy that would likely occur across winter, spring, and summer gaps in communication instruction following a one-course sequence. Yet, a helpful line of future research might use MLM to consider differences in mean self-efficacy outcomes for BCC programs that implement two-semester versus single-semester course modalities. If colleges aim to ready their students for future careers, additional attention must be paid to fostering students' communication skills and maintaining them over time.

Finally, several issues related to measurement are present in the research, specifically the use of items from Ellis' (1995) Self-Perceived Public Speaking Competence scale. Previous research indicates that the 4 items taken from that measure for this study are unidimensional (i.e., Stewart et al., 2017; Stewart et al., 2019); however, the fit statistics obtained through the invariance testing indicate poor fit at each wave. Moreover, the present results echo previous research by demonstrating that the measure is unreliable (ω at each wave < 0.65). That is, there is increased error in the measurement of the construct. This inflates the possibility of not observing a significant result when it truly exists (i.e., Type II error) and can impact overall fit statistics during structural equation modeling (i.e., invariance analyses; Yetkiner & Thompson, 2010). Given the inconsistent performance of the measure in this study and others, we encourage researchers to avoid using the measure in any capacity, much less as a proxy for students' self-efficacy.

Indeed, although the scale is similar in nature to the theoretical conceptualization of self-efficacy, it lacks isomorphism with the proposed construct, particularly in the use of *can-do* statements related to specific behaviors (Usher, 2015). The items taken from the scale also fail to capture *content* as an important component of effective public speaking, focusing primarily on student skills related to *delivery*. Researchers have made attempts to specifically evaluate public speaking self-efficacy (e.g., Warren, 2011), but no validated measure currently exists to accurately operationalize the construct. Until this measure exists, studies of public speaking self-efficacy should be interpreted with care.

Conclusion

The BCC is constantly evolving (Valenzano et al., 2014). As it begins to take on new forms and address a variety of skills that students and employers find useful, scholars should respond by considering the use of self-efficacy to approximate actual capabilities. A greater focus on the differences in students' unique educational experiences, as well as the use of tools that model learning as a result of change over time, should help administrators adapt to students' needs and create a more inclusive educational experience in one of the most important and foundational courses to the communication discipline. Our call for increased attention to the way identities and individual differences influence assessment is not new, and we hope that researchers will consider the questions raised in this manuscript with care.

Bibliography

Ashby-King, D. T. (2021). More than just a variable: COVID-19 and the call to complicate communication education research. *Communication Education, 70*(2), 205-207. https://doi.org/10.1080/03634523.2020.1857417

Autman, H., & Kelly, S. (2017). Reexamining the writing apprehension measure. *Business and Professional Communication Quarterly*, *80*(4), 516-529. https://doi.org/10.1177%2F2329490617691968

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

Bauer-Wolf, J. (2019, January 17). *Survey: Employers want 'soft skills' from graduates.* Inside Higher Ed. https://www.insidehighered.com/quicktakes/2019/01/17/survey-employers-want-soft-skills-graduates

Bliese, P. D., & Hanges, P. J. (2004). Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods, 7*(4), 400-417. https://doi.org/10.1177%2F1094428104268542

Bodie, G. D. (2010). A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. *Communication Education, 59*(1), 70-105. https://doi.org/10.1080/03634520903443849

Broeckelman-Post, M. A., & Pyle, A. S. (2017). Public speaking versus hybrid introductory communication courses: Exploring four outcomes. *Communication Education, 66*(2), 210-228. https://doi.org/10.1080/03634523.2016.1259485

Broeckelman-Post, M. A., Hawkins, K. E. H., Murphy, J., Otusanya, A., & Kueppers, G. (2020). The impact of gender and introductory communication course type on public speaking performance. *Communication Teacher, 34*(1), 53-67. https://doi.org/10.1080/17404622.2019.1593478

Burns, E. C., Martin, A. J., Kennett, R. K., Pearson, J., & Munro-Smith, V. (2021). Optimizing science self-efficacy: A multilevel examination of the moderating

effects of anxiety on the relationship between self-efficacy and achievement in science. *Contemporary Educational Psychology, 64*, Article 101937. https://doi.org/10.1016/j.cedpsych.2020.101937

Byars-Winston, A., Diestelmann, J., Savoy, J. N., & Hoyt, W. T. (2017). Unique effects and moderators of effects of sources on self-efficacy: A model-based meta-analysis. *Journal of Counseling Psychology*, *64*(6), 645 –658. https://doi.org/10.1037/cou0000219

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466. https://doi.org/10.1037/0033-2909.105.3.456

Byrne, M., Flood, B., & Griffin, J. (2014). Measuring the academic self-efficacy of first-year accounting students. *Accounting Education, 23*(5), 407-423. https://doi.org/10.1080/09639284.2014.931240

Campbell, K. K. (1989). *Man cannot speak for her, volume I: A critical study of early feminist rhetoric.* Greenwood Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling: A multidisciplinary journal, 14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling: A multidisciplinary journal*, *9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Chesebro, J. W., McCroskey, J. C., Atwater, D. F., Bahrenfuss, R. M., Cawelti, G., Gaudino, J. L., & Hodges, H. (1992). Communication apprehension and self-perceived communication competence of at-risk students. *Communication Education*, *41*(4), 345-360. https://doi.org/10.1080/03634529209378897

Cleary, L. M. (1996). "I think I know what my teachers want now": Gender and writing motivation. *The English Journal, 85*(1), 50-57. https://doi.org/10.2307/821123

Daly, J. A. (1978). Writing apprehension and writing competency. *The Journal of Educational Research, 72*(1), 10-14. https://doi.org/10.1080/00220671.1978.10885110

Dedrick, R. F., Ferron, J. M., Hess, M. R., Hogarty, K. Y., Kromrey, J. D., Lang, T. R., Niles, J. D., & Lee, R. S. (2009). Multilevel modeling: A review of methodological issues and applications. *Review of Educational Research*, *79*(1), 69-102. https://doi.org/10.3102/0034654308325581

Dresel, M., & Rindermann, H. (2011). Counseling university instructors based on student evaluations of their teaching effectiveness: A multilevel test of its effectiveness under consideration of bias and unfairness variables. *Research in Higher Education, 52*(7), 717–737. https://doi.org/10.1007/s11162-011-9214-7

Dwyer, K. K., & Fus, D. A. (2002). Perceptions of communication competence, self-efficacy, and trait communication apprehension: Is there an impact on basic course success? *Communication Research Reports, 19*(1), 29-37. https://doi.org/10.1080/08824090209384829

Ellis, K. (1995). Apprehension, self-perceived competency, and teacher immediacy in the laboratory-supported public speaking course. *Communication Education, 44*(1), 64-78. https://doi.org/10.1080/03634529509378998

Faulkner, S. L., Watson, W. K., Pollino, M. A., & Shetterly, J. R. (2021). "Treat me like a person, rather than another number": University student perceptions of inclusive classroom practices. *Communication Education, 70*(1), 92-111. https://doi.org/10.1080/03634523.2020.1812680

Frey, T. K. & Lane, D. R. (2021). CAT in the classroom: A multilevel analysis of students' experiences with instructor nonaccommodation. *Communication Education, 70*(3), 1-24. https://doi.org/10.1080/03634523.2021.1903521

Frey, T. K., & Vallade, J. (2018). Assessing students' perceptions of writing and public speaking self-efficacy in a composition and communication course. *Journal of Communication Pedagogy, 1*, 27-39. https://doi.org/10.31446/JCP.2018.08.

Frisby, B. N., Kaufmann, R. M., Vallade, J. I., Frey, T. K., & Martin, J. C. (2020). Using virtual reality for speech rehearsals: An innovative instructor approach to enhance student public speaking efficacy. *Basic Communication Course Annual, 32*, 59-78. https://ecommons.udayton.edu/bcca/vol32/iss1/6/

Goodboy, A. K., Bolkan, S., Kromka, S. M., Knoster, K. C. (2021). Instructional dissent over the course of the semester. *Communication Education*, 1-18. Advance online publication. https://doi.org/10.1080/03634523.2021.1925718

Hampsten, K. (2021). Embracing discomfort and resisting a return to "the good old days": A call to communication educators. *Communication Education, 70*(2), 208-210. https://doi.org/10.1080/03634523.2020.1857413

Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research, 32*, 385-410. https://doi.org/10.1111/j.1468-2958.2006.00281.x

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But... *Communication Methods and Measures, 14*(1), 1-24. https://doi.org/10.1080/19312458.2020.1718629

Hox, J. (2002). *Multilevel analysis techniques and applications*. Lawrence Erlbaum Associates Publishers.

Huang, C. (2012). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education, 28*, 1-35. https://doi.org/10.1007/s10212-011-0097-y

Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education*, *84*(1), 175-196. https://doi.org/10.1080/00220973.2014.952397

Huang, F. L. (2018). Multilevel modeling myths. *School Psychology Quarterly*, *33*(3), 492–499. https://doi.org/10.1037/spq0000272

Hunter, K. M., Westwick, J. N., & Haleta, L. L. (2014). Assessing success: The impacts of a fundamentals of speech course on decreasing public speaking anxiety. *Communication Education, 63*(2), 124-135. https://doi.org/10.1080/03634523.2013.875213

Kreft, I. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30*(1), 1-21. https://doi.org/10.1207/s15327906mbr3001_1

Lane, D. R. (2017). Raising new questions and restoring our focus on authentic student learning. *Communication Education, 66*(1), 120-122. https://doi.org/10.1080/03634523.2016.1243794

LeFebvre, L., LeFebvre, L. E., & Allen, M. (2018). Training the butterflies to fly in formation: Cataloguing student fears about public speaking. *Communication Education, 67*(3), 348-362. https://doi.org/10.1080/03634523.2018.1468915

LeFebvre, L., LeFebvre, L. E., Allen, M., Buckner, M. M., & Griffin, D. (2020). Metamorphosis of public speaking anxiety: Student fear transformation throughout the introductory communication course. *Communication Studies, 71*(1), 98-111. https://doi.org/10.1080/10510974.2019.1661867

Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford Press.

Ma, X., Ma, L., & Bradley, K. (2008). Using multilevel modeling to investigate school effects. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 59-110). Information Age.

Mamaril, N. J. A. (2014). *Measuring undergraduate students' engineering self-efficacy: A scale validation study*. [Doctoral dissertation, University of Kentucky]. https://uknowledge.uky.edu/edp_etds/19

McCoach, D. B., & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology*, *1*, 1-17. https://doi.org/10.3389/fpsyg.2010.00017

McCroskey, J. C. (1982). Oral communication apprehension: A reconceptualization. In M. Burgoon (Ed.), *Communication yearbook 6*, (pp. 136-170). Sage.

McKenna-Buchanan, T., Munz, S., Wright, A., & Williams, J. (2020). The importance of the basic communication course in the first-year experience: Implications for retention. *Basic Communication Course Annual, 32*(1), 148-170. https://ecommons.udayton.edu/bcca/vol32/iss1/10

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568-592. https://doi.org/10.1037/0021-9010.93.3.568

Morreale, S. P., Shockley-Zalabak, P. S., Gaddis, B., Thorpe, J., Staley, C. M., Allgood, E. (2021). A 14-year empirical analysis of undergraduates' pre- and post-test scores in three introductory communication courses: Lessons learned for pedagogy and assessment. *Basic Communication Course Annual, 33*, 128-159. https://ecommons.udayton.edu/bcca/vol33/iss1/9

Mottet, T. P., & Richmond, V. P. (1998). Newer is not necessarily better a reexamination of affective learning measurement. *Communication Research Reports*, *15*(4), 370-378. https://doi.org/10.1080/08824099809362136

National Center for Education Statistics. (2019). *Enrollment.* https://nces.ed.gov/fastfacts/display.asp?id=98

National Communication Association. (2020). *Basic course & general education*. https://www.natcom.org/academic-professional-resources/teaching-and-learning/basic-course-general-education

Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, *2*(2), 842-860. https://doi.org/10.1111/j.1751-9004.2007.00059.x

Nordin, K., & Broeckelman-Post, M. A. (2019). Can I get better? Exploring mindset theory in the introductory communication course. *Communication Education, 68*(1), 44-60. https://doi.org/10.1080/03634523.2018.1538522

Nordin, K., & Broeckelman-Post, M. A. (2020). Surviving or thriving? Demographics differences in mindset across the introductory communication course. *Communication Education, 69*(1), 85-104. https://doi.org/10.1080/03634523.2019.1679379

Nyquist, J. L., & Booth, J. L. (1977). Instructional communication: A basic course for teachers. *Communication Education, 26*(1), 13-26. https://doi.org/10.1080/03634527709378195

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly, 19*(2), 139-158. https://doi.org/10.1080/10573560308222

Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English, 42*(1), 104-120. https://www.jstor.org/stable/40171749

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*(1), 71-90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.

Raudenbush, S. W., Bryk, A. S, Cheong, Y.F., & Congdon, R. (2019). *HLM 8 for Windows* [Computer software]. Skokie, IL: Scientific Software International, Inc. https://ssicentral.com

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. http://dx.doi.org/10.1037/a0026838

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. http://dx.doi.org/10.18637/jss.v048.i02

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. https://doi.org/10.1007/BF02296192

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin, 143*(6), 565–600. https://doi.org/10.1037/bul0000098

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*(1), 91-100. https://doi.org/10.1037/0022-0663.81.1.91

Simmons, J., & Wahl, S. T. (2016). Rethinking inclusion and diversity in communication education research. *Communication Education, 65*(2), 232–235. https://doi.org 10.1080/03634523.2015. 1098711

Strawser, M. G., Housley Gaffney, A. L., DeVito, A., Kercsmar, S. E., & Pennell, M. (2017). A blended basic course examination of communication apprehension and self-efficacy: A comparative analysis. *Basic Communication Course Annual, 29*, 53-73. https://ecommons.udayton.edu/bcca/vol29/iss1/6/

Stewart, C. O., McConnell III, J. R., Stallings, L. A., & Roscoe, R. D. (2017). An initial exploration of students' mindsets, attitudes, and beliefs about public speaking. *Communication Research Reports*, *34*(2), 180-185. https://doi.org/10.1080/08824096.2016.1270821

Stewart, C. O., McConnell III, J. R., Stallings, L. A., & Roscoe, R. D. (2019). Growth mindset: Associations with apprehension, self-perceived competence, and beliefs about public speaking. *Basic Communication Course Annual, 31*(1), 44-69. https://ecommons.udayton.edu/bcca/vol31/iss1/6

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th edition). Pearson.

Usher, E. L. (2015). Personal capability beliefs. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (3rd ed., pp. 146-159). Taylor & Francis. https://doi.org/10.4324/9781315688244

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751-796. https://doi.org/10.3102/0034654308321456

Valenzano III, J. M., Wallace, S. P., & Morreale, S. P. (2014). Consistency and change: The (r) evolution of the basic communication course. *Communication Education*, *63*(4), 355-365. https://doi.org/10.1080/03634523.2014.911928

van Buuren, S. (2011). Multiple imputation of multilevel data. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of Advanced Multilevel Analysis* (pp. 173-196). Routledge.

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Wallace, S. P. (2014). Student learning outcomes: Primary drivers of course design. *Basic Communication Course Annual, 26*, 1-8. https://ecommons.udayton.edu/bcca/vol26/iss1/6

Warren, J. L. (2011). *The relationship between service learning and public speaking self-efficacy: Toward engaging today's undergraduates*. [Doctoral dissertation, University of Kentucky]. https://uknowledge.uky.edu/comm_etds/1/

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10-18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Wilkins, J. L., & Ma, X. (2002). Predicting student growth in mathematical content knowledge. *The Journal of Educational Research, 95*(5), 288-298. https://doi.org/10.1080/00220670209596602

Yetkiner, Z. E., & Thompson, B. (2010). Demonstration of how score reliability is integrated into SEM and how reliability affects all statistical analyses. *Multiple Linear Regression Viewpoints, 36(*2), 1-13.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663-676. https://doi.org/10.3102%2F00028312029003663

Table 1. Longitudinal Measurement Invariance Model Fit Statistics for WSE and SPPSC across Waves 1-4.

Model	χ2 (*df*)	CFI	TLI	SRMR	RMSEA (90% CI)	MC	$\Delta\chi^2$ (*df*)	ΔCFI	ΔTLI	ΔRMSEA
WSE										
M1: Wave 1	21.478 (14)	.986	.979	.030	.039 (.000, .065)					
M2: Wave 2	39.418 (14)***	.963	.945	.038	.072 (.052, .093)					
M3: Wave 3	24.451 (14)*	.986	.980	.027	.046 (.022, .069)					
M4: Wave 4	21.292 (14)	.991	.986	.025	.039 (.008, .061)					
M5: Configural	420.639 (303)***	.974	.968	.058	.033 (.026, .040)					
M6: Metric	452.526 (321)***	.971	.966	.086	.034 (.028, .041)	M5	32.908 (18)*	-0.003	-0.002	0.001
M6b: Partial Metric	443.967 (320)***	.973	.968	.073	.033 (.026, .040)	M5	23.285 (17)	-0.001	0.000	0.000
M7a: Scalar	477.036 (337)***	.969	.967	.076	.035 (.028, .041)	M6b	35.843 (17)**	-0.004	-0.001	0.002
M7b: Partial Scalar	464.679 (335)***	.972	.968	.074	.033 (.027, .040)	M6b	20.683 (15)	-0.001	0.000	0.000
SPPSC										
M8: Wave 1	8.350 (2)*	.938	.814	.036	.095 (.044, .154)					
M9: Wave 2	10.830 (2)**	.943	.829	.038	.112 (.061, .171)					
M10: Wave 3	18.157 (2)***	.872	.615	.053	.152 (.105, .204)					
M11: Wave 4	23.176 (2)***	.863	.588	.069	.174 (.129, .224)					
M12: Configural	156.606(75)***	.932	.891	.061	.056 (.044, .067)					
M13: Metric	162.546 (84)***	.934	.906	.067	.052 (.041, .063)	M12	5.323 (9)	0.002	0.015	-0.004
M14a: Scalar	194.426 (93)***	.915	.891	.076	.056 (.046, .066)	M13	33.953 (9)***	-0.019	-0.015	0.004
M14b: Partial Scalar	174.870 (91)***	.930	.908	.073	.051 (.041, .062)	M13	12.155 (7)	-0.004	0.002	-0.005

Note. *p < .05, **p < .01, ***p < .001; χ^2 = scaled chi-square test statistic; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval around RMSEA; MC = models compared; $\Delta\chi^2$ = change in scaled χ^2 relative to the prior model; ΔCFI = change in comparative fit index relative to the prior model; ΔTLI = change in Tucker-Lewis index relative to the prior model; ΔRMSEA = change in root mean square error of approximation.

Figure 1. Mean Growth Plots for WSE and SPPSC Across Four Waves

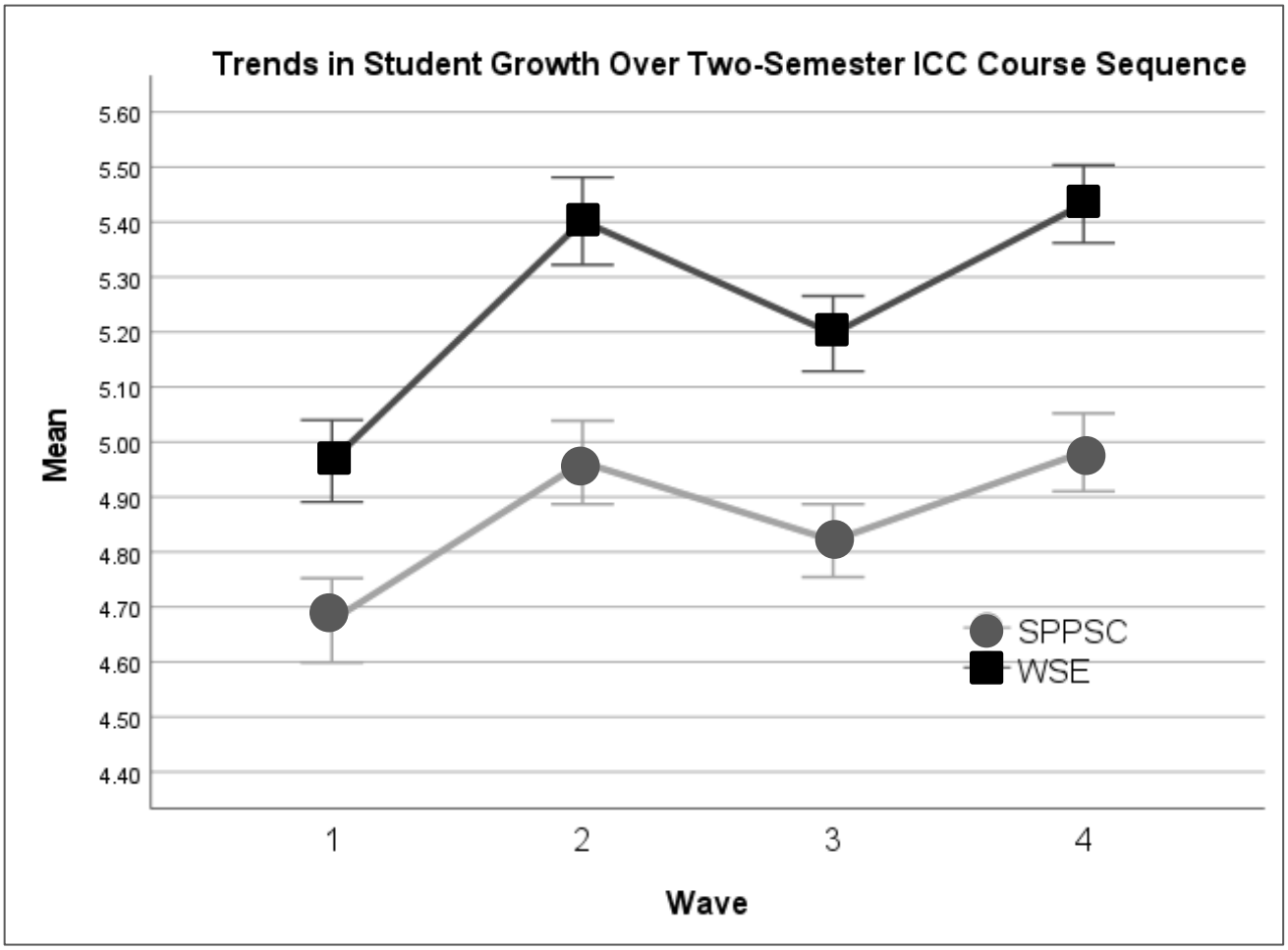

Table 2. Unconditional Linear Growth Models

	WSE		SPPSC	
Fixed effect	Coefficient	SE	Coefficient	SE
Initial status	5.00	.03***	4.69	.03***
Growth	0.15	.01***	0.10	.01***
Random Effect	Variance Component			
Level 1				
Within student	0.36		0.34	
Level 2				
Initial Status	0.45***		0.48***	
Growth	0.00		0.03***	

Note. SE = standard error.
*$p < .05$, **$p < .01$, ***$p < .001$

Table 3. Linear Growth Models with Time-Varying Covariate (TVC)

	WSE		SPPSC	
Fixed effect	Coefficient	SE	Coefficient	SE
Initial status	4.98	.03***	4.68	.03***
Growth	0.32	.02***	0.22	.02***
Winter Break (TVC)	-0.47	.05***	-0.34	.05***
Random Effect	Variance Component			
Level 1				
Within student	0.34		0.33	
Level 2				
Initial Status	0.46***		0.49***	
Growth	0.01*		0.03***	

Note. SE = standard error.
*$p < .05$, **$p < .01$, ***$p < .001$

Table 4. Final Models Predicting Individual Growth and Initial Status for WSE and SPPSC

	WSE		SPPSC	
Variable	Effect	SE	Effect	SE
Effects of between-student variables on initial status				
Intercept	4.90	.04***	4.56	.04***
Sex	0.17	.05***	0.23	.05***
Writing Affect	0.12	.03***	---	---
Writing Apprehension	-0.36	.03***	---	---
Public Speaking Apprehension	---	---	-0.45	.04***
Public Speaking Affect	---	---	0.14	.03***
Effects of between-student variables on rate of growth				
Intercept	0.32	.02***	0.22	.02***
Sex	0.04	.02	0.00	.03
Writing Affect	-0.03	.01	---	---
Writing Apprehension	0.06	.01***	---	---
Public Speaking Apprehension	---	---	0.06	.02***
Public Speaking Affect	---	---	0.00	.02
Effect of time-varying covariate				
Winter Break	-0.50	.05***	-0.36	.05***

Note. SE = standard error.
*$p < .05$, **$p < .01$, ***$p < .001$