Classifying the quality of robotics-enhanced lesson plans using motivation variables, word count, and sentiment analysis of reflections

Brian R. Belland | bbelland@psu.edu

ChanMin Kim | cmk604@psu.edu

Anna Y. Zhang | ybz5148@psu.edu

Eunseo Lee | eul374@psu.edu

Emre Dinc | ezd287@psu.edu

The Pennsylvania State University

106 CEDAR

University Park, PA 16802 USA

This is an author-created version of the article published by Elsevier in Contemporary Educational Psychology:

Belland, B. R., Kim, C., *Zhang, A., *Lee, E., & *Dinç, E. (2022). Classifying the quality of robotics-enhanced lesson plans using motivation variables, word count, and sentiment analysis of reflections. *Contemporary Educational Psychology, 69*, 1-11.

https://doi.org/10.1016/j.cedpsych.2022.102058

Introduction

Identified as the central context within which culture evolves and is appropriated within early childhood years, play has long been key to early childhood education (ECE) (Brooker et al., 2014; Vygotsky, 1978). Though play, children learn cultural norms and traditions, and negotiate their place within their community (Vygotsky, 1978). But this does not mean that ECE teachers sit back as young children play. Rather, ECE teachers arrange types and sequences of play experiences that can help young children gain cognitive, socio-emotional, and communicative competencies (Aras, 2016; Ashiabi, 2007). Often, play may only involve children and their own imagination. However, play can also involve manipulables, such as educational robots (Bers, 2008). Educational robots can serve as co-players in children's dramatic play. Using robots in such a way can improve (a) executive function (Di Lieto et al., 2017), (b) problem solving abilities (Toh et al., 2016), (c) interest in engineering and technology (Sullivan & Bers, 2019), and (d) spatial awareness (Torres et al., 2018) among early childhood learners. But facilitating this is no small feat, and thus careful teacher preparation is needed. Specifically, prospective teachers need to learn to engage in programming, work with robots, and integrate such into well-designed lesson plans. In this paper, we investigate classifiers of the quality with which prospective, early childhood teachers integrated robots and block-based coding in lesson plans. In the next section we review the literatures on how prospective teachers learn to design lesson plans, teacher identity, teacher motivation, and the use of robots in early childhood education. Next, we state our research questions, *How can prospective teachers' lesson plan quality be classified using motivation and process variables*, *which prospective teachers' lesson plans are misclassified*, *and why*, and *how do motivation and process variables predict prospective teachers' lesson plan quality*? Then, we advance our methodology using (a) discriminant analysis and ordinal logistic regression to classify and predict lesson plan quality using process and motivation variables, and (b) support vector machines to identify misclassified lesson plans. Next, results are presented, and after that discussed in light of the literature, including limitations and conclusions.

Conceptual Framework: Teachers as Learners for Lesson Design with Robots

The present study's conceptual framework is grounded in two fundamental perspectives on teacher learning: *teachers as learners* (e.g., Kwo, 2010; Middlewood et al., 2005; Plate & Peacock, 2021) and *teachers as designers* (e.g., Kali et al., 2015; Laurillard, 2012; McKenney et al., 2015). The teachers as learners perspective considers the ongoing and deep learning required to become and be a teacher (Näykki et al., 2021a; Shulman & Sherin, 2004). There is much to learn as a prospective teacher, including, but not limited to, strategies to (a) teach different content, (b) maintain classroom discipline, (c) sustain student motivation, and (d) write plans for effective and engaging lessons. Teacher learning then needs to continue as a practicing teacher. But just as prospective teachers need to learn these things, their own motivation and learning processes are critical to how well they proceed in their journeys to become teachers. It is critical to carefully examine the process by which prospective teachers learn, and indicators that can explain the quality with which skills are learned. Inasmuch as prospective teachers are learners, both prospective and practicing teachers are designers (Asensio-Pérez et al., 2017; Kali et al., 2015). Teachers not only design lessons but also engage in an engineering-style design process, which requires that teachers produce an optimal solution given the constraints of materials, resources, and time. For example, Kali et al. (2015) referred to teaching as a design science and argued for the critical role of design in teacher learning of technology and with technology. The present study reports empirical data on how prospective ECE teachers learned to design technology (robots) and design lessons using their designed robots.

The present study advances the relevant literatures by integrating two perspectives into early childhood teacher education for CS learning and robotics inclusion. Such integrated perspectives enabled three research-based foci that guided the study. First, positioning teachers as learners should begin from teacher preparation programs. Being teachers as learners involves continuous, self-activated learning in response to emergent needs (Middlewood et al., 2005). This is unlikely to occur without practice, and practice is needed starting in prospective teacher years, and continuing throughout teaching careers. In a teachers-as-learners community designed for self-regulated learning of both prospective and practicing teachers, prospective teachers reported the importance of motivation in engagement with continuous teacher learning (Näykki et al., 2021b). Second, learning to work with new but nonmandatory content and methods is also critical in becoming teachers as learners who respond to emergent needs. While robot inclusion is not mandatory in ECE, it responds to emergent needs of young children to explore STEM pathways (Çetin & Demircan, 2020) and engage with a world in which technology is ever more present and central to life (Zviel-Girshin et al., 2020). While some robots can be controlled using buttons, many are programmed using block-based programming, defined as visual representations of programming functions, loops, and variables (Grover & Basu, 2017). Key to promoting the use of robots in ECE is inviting and supporting ECE teachers to learn to code (Jaipal-Jamani & Angeli, 2018; Kidd et al., 2020). Thus, ECE teacher learning of robotics is part of their becoming teachers as learners. Such learning is also possible through design. Design has been used in teacher education (e.g., Koehler et al., 2007; C.-J. Lee & Kim, 2014). The literature on teachers as designers highlights the importance of teacher learning through design (Cober et al., 2015; Kali et al., 2015; Matuk et al., 2015; Svihla et al., 2015; Voogt et al., 2015). Design thinking is emphasized in teacher education especially on technology use (C.-C. Tsai & Chai, 2012). Learning to "think like designers" can enable learners to overcome difficulties and solve real-world problems (Razzouk & Shute, 2012, p. 343). Last, the process of design and learning through reflection is crucial because reflection is required in design (Røise et al., 2014) and central to learning to teach (Hofer, 2017; Martinez et al., 2019). Each of these research-based foci is described further in the next three sections: motivation of teachers as learners, robot inclusion by teachers as designers, and process of learning and design through reflection.

Motivation of Teachers as Learners

Motivation to teach can be considered from a variety of perspectives that consider the role of goal orientations (Parker et al., 2012; Runhaar et al., 2010), expectancy for success (Liang & Tsai, 2008; Sak, 2015), beliefs in the value of an outcome (Foley, 2011), identification with the subject being taught (Akkerman & Meijer, 2011), and perceptions of belongingness (Arndt, 2018) and autonomy (Skaalvik & Skaalvik, 2014). Individuals can hold mastery goal orientations, in which they strive to master content when engaging in learning tasks, or performance goal orientations, in which they strive to perform better than others (Pintrich, 2000). Expectancy for success refers to one's perception of the probability of success doing different activities, such as using the Internet (Liang & Tsai, 2008), and integrating technology (Y. Lee & Lee, 2014). Having low expectancy for success in teaching science may lead an ECE teacher to avoid teaching science as much as possible, or to not exude confidence when teaching science (Mintzes et al., 2013). This, in turn, can lead to poor science learning outcomes among ECE students (Mintzes et al., 2013), which is especially problematic because early success sets the stage for continued refinement and deeper learning (Bransford & Schwartz, 1999). If teachers have insufficient belief in the value of teaching in a particular way (e.g., with robots), they will likely not invest considerable time designing a lesson to teach in that way (Foley, 2011).

The ways in which teachers identify with particular people, perspectives, and content areas also have an outsize influence on the way that they teach (Akkerman & Meijer, 2011). Teachers' identities incorporate their views on the role of teachers as authority figures and either the source of knowledge or facilitator of construction of knowledge among students (Friesen & Besley, 2013). Furthermore, teachers' identities influence the way that they learn, including how to plan lessons (Beijaard, 2017). For example, teachers who see their identity as that of an expert who needs to convey that expertise to students will likely design lecture-based lessons. Still, teacher identity is not a monolithic construct; rather, teacher identity is shaped by and evolves through an accumulation of experiences acquired before, during, and after teacher education (Akkerman & Meijer, 2011). One cannot understand how a teacher teaches without also understanding how she frames her own identity (Olsen, 2008). Within the context of learning to teach computer science content, it is important to address teachers' identification with mathematics, engineering, and computer science (Ni & Guzdial, 2012).

The present study was expected to produce knowledge about critical motivation, among these possible motivational factors, of *teachers as learners* related to their teaching with robots. It was expected that the unique context of the present study, with participants who were neither teachers yet (considering they were prospective teachers) nor learners alone (considering they taught two days per week in preschools) but becoming teachers as learners, would produce knowledge about critical motivation of *teachers as learners* related to their teaching with robots.

Robot Inclusion by Teachers as Designers

As discussed earlier, learning to work with new but nonmandatory content and methods is critical part of becoming teachers as learners. Such learning can be achieved through design beyond lesson planning alone. This is partly because new content or method *learning* is necessary to teach the new content or with the new method. While design-based learning has been widely used in STEM education for a variety of students (e.g., Doppelt, 2009), design is used also for *teacher* learning of STEM in teacher education (e.g., Kim et al., 2015). But design in teacher learning contexts does not stop at learning of new STEM content and methods. Design goes into planning to teach with what is designed. Learning through design and design for teaching is crucial to becoming teachers as designers (Kali et al., 2015).

One can determine the preparedness of teachers as designers to integrate robotics effectively into teaching by rating their lesson plans in which the robots that *they designed* are used. Prospective teachers need to learn to design coherent lesson plans that help diverse children meet learning objectives, defined as plans for teaching that carefully document learning objectives, learner characteristics (e.g., prior knowledge and interests), and lesson activities and assessment (Pashler et al., 2007; Sias et al., 2017). Lesson plans are used to guide teaching, but the best lesson plans are sufficiently detailed to include specific teaching strategies, yet flexible enough to allow for reaction to student struggles (Y.-A. Lee & Takahashi, 2011; Sias et al., 2017). A good lesson plan should be seen as a framework to guide teaching, rather than a sequence of activities to be followed in a lock-step manner (Ding & Carlson, 2013; John, 2006).

Lesson planning skills are often considered so critical that prospective and practicing teacher quality is often assessed by rating lesson plan quality (Backfisch et al., 2020; Musselwhite & Wesolowski, 2018; Ozogul et al., 2008; P.-S. Tsai & Tsai, 2019). Rating lesson plans can provide insight into teachers' intentions for teaching, including their intentions to meaningfully integrate technology (Janssen & Lazonder, 2016; Pringle et al., 2015; P.-S. Tsai &

Tsai, 2019) and their use of constructivist approaches (Backfisch et al., 2020; König et al., 2020; Milner et al., 2011). Rating criteria often focuses on such areas as instructional planning (e.g., specifying and responding to students' prior knowledge), instructional strategies (e.g., providing guidance), and assessment (e.g., quality with which lesson provides for assessing students).

Process of Learning and Design through Reflection

Critical reflection on learning processes is central to learning to teach (Hofer, 2017; Martinez et al., 2019) and in design (Røise et al., 2014). Such reflection can take many forms, including retrospective reflection and just-in time reflection (Hofer, 2017). In retrospective reflection, teachers think back to events and processes that took place in the past, and about what such events and processes mean to their evolving teaching skills and inclinations. Just in time reflection involves thinking about events and processes as they are happening, and the implications of such for the teacher's evolving teaching skills and inclinations. One way to judge depth of reflection is to apply a rubric to each reflection entry (Molee et al., 2011). But when prospective teachers engage in multiple reflections, and the goal is to count the number of words used in the reflection, using a rubric to assess each reflection can be cumbersome. Furthermore, using a rubric in such a way cannot provide the type of just-in-time information needed to intervene among prospective teachers who are on track to produce a low-quality lesson plan. An alternative that provides reasonably valid information about depth of reflection is word count (Davis, 2006; Wulff et al., 2021). Notably, longer word counts can evidence deeper reflection.

One can also assess the sentiment reflected in the reflection posts (Liu, 2012; Wen et al., 2014). This process involves feeding the writing into a computer algorithm that assesses the writing on a scale of -1 to 1, where -1 indicates the most negatively valanced writing, while 1 represents the most positively valanced writing. Sentiment analysis has been applied in educational research, often evaluating open ended course evaluation comments (Dolianiti et al., 2019). In addition, sentiment analysis can provide a relatively objective metric related to motivation. Motivation is often measured using self-report surveys and qualitative analyses of observation and prompted, retrospective interview data (Fulmer & Frijters, 2009). Limitations to self-report data include presentation bias (Kopcha, 2007) and the extent to which children can be expected to understand the constructs being assessed (Fulmer & Frijters, 2009). Still, when assessing such constructs such as expectancy for success and belongingness among adults, presentation bias and ability to understand the constructs seem to be of little concern. Still, it is wise to consider other methodologies with which one can gather additional data on motivation.

Research Questions and Hypotheses

While lesson plan quality ratings are a good indicator of the preparedness of prospective teachers as designers to integrate robotics, it is also clear that such ratings often come too late to intervene to help prospective teachers improve. Notably, identifying prospective teachers who are on track to successfully design robots and integrate the robots that they design into their lesson plans is critical so that there is enough time to provide additional support to those who are not on track. One can do so using discriminant analysis to assign prospective teachers to classes corresponding to different lesson plan qualities (Lachenbruch & Goldstein, 1979). Lesson plans are often produced at the end of a unit or an entire course, at which point it is often too late for just-in-time scaffolding. Discriminant analysis produces coefficients of linear discriminants that can be used to predict membership in low, average, or high-quality lesson plan classes.

As discussed in the conceptual framework of the present study, relevant inputs to such linear discriminant models include prospective teachers' motivation to teach with robots, their identity, and the depth with which they reflect on their processes of learning to teach with robots and teaching in field experience. Hence, this study seeks to answer the following research questions:

1. How can prospective teachers' lesson plan quality be classified using motivation and process variables?
2. How do motivation and process variables predict prospective teachers' lesson plan quality?

Relative to research question 1, we hypothesized that at least one of the following variables - perception of mathematics, CS and engineering emphasis in STEM career, views of coding, perceptions of computer science and technology, science interest, performance goal orientation, perceptions of English, mastery goal orientation, perceptions of computer science, STEM emotions, and word count - can be used to classify lesson plan quality.

Relative to research question 2, we hypothesized that at least one of the following variables - perception of mathematics, CS and engineering emphasis in STEM career, views of coding, perceptions of computer science and technology, science interest, performance goal orientation, perceptions of English, mastery goal orientation, perceptions of computer science, STEM emotions, and word count - can be used to predict lesson plan quality.

Method

Participants and Setting

The research setting was an early childhood education course on integrating performing and visual arts to enhance communication, inquiry, and engagement in P-5 education, which was offered by a large public university in the United States. The course also included a field experience component at local preschools. Robotics activities took place in three class sessions that covered robotics in early childhood education, STEM education, robot programming, and sample lessons. Prospective teachers in the course programmed robots and designed preschool lessons using the robots. Forty-six prospective teachers participated. Seven participants were excluded because they did not complete the culminating lesson plan. All participants were female; 32 (82.05%) were white, 3 (7.7%) were Hispanic, 2 (5.13%) were Asian, 1 (2.56%) was black, and 1 (2.56%) was multi-racial. The age range was 19-22 (M = 20.31, SD = 0.92). On the presurvey, 22 participants (56.41%) self-reported having no computer programming knowledge, 10 (25.64%) self-reported having low computer programming knowledge, and 7 (17.95%) self-reported having intermediate computer programming knowledge. Thirty-two (82.05%) participants had no experience with educational robot programming, while 7 (17.95%) had educational robot programming experience.

Materials

Ozobot and OzoBlockly

Ozobot Bit is a small robot that can (a) create diverse movements and lights, (b) sense lines and colors, and (c) be integrated into problem-solving tasks and school curriculums (Hunsaker, 2018). Ozobot Bit follows block-based code created using the OzoBlockly platform.

Lesson Plan Template

Participants were required to design a lesson using robots that they programmed, and implement the lesson in their field experience class. A lesson plan template included essential lesson plan components. The template was organized into the sections of lesson goals, objectives, considerations (e.g., materials, prior knowledge), and details of class activities.

Data Collection

Presurvey

We conducted principal components analysis with varimax rotation using data from 8 class sections (N = 176) that participated in the overall project. The Kaiser-Meyer-Olkin value (= 0.801) indicated that the sample size was adequate (Kaiser, 1974). The presurvey contained 100 items and covered ten latent variables of motivation. Parallel analysis, in addition to the criterion of eigenvalues, indicated a ten-factor solution (see Section A in the supplementary material) explaining 62.38% of the variance.

Items loading on Factor 1, perceptions of mathematics, addressed participants' interest, self-efficacy, and achievement emotions in mathematics. This factor included 16 items with factor loadings larger than 0.414 and for which Cronbach's α was 0.952 (note: all cronbach's α values presented here were calculated from the current dataset). Factor 2, computer science and engineering emphasis in STEM career, covered interest in computer science and engineering-related careers. Fifteen items with factor loadings larger than 0.427 loaded on this factor and Cronbach's α was 0.946. Factor 3, views of coding, covered students' attitudes toward coding knowledge and skills. Twelve items with factor loadings larger than 0.394 were included and the Cronbach's α for this factor was 0.902. Factor 4, perceptions of computer science and technology, is associated with interest and self-efficacy in computer science, technology, and engineering. Eleven items with factor loadings larger than 0.365 loaded on this factor and Cronbach's α was 0.882. Factor 5, science interest, covered interest in science. Ten items with factor loadings larger than 0.400 were included in this factor and the Cronbach's α was 0.887. Factor 6, performance goal orientation, covered performance-approach and performance-avoid goal orientations. Eleven items with factor loadings larger than 0.355 loaded on this factor and Cronbach's α was 0.870. Factor 7, perceptions of English, covered interest and self-efficacy in English. Seven items with factor loadings larger than 0.596 were included in this factor and the Cronbach's α was 0.920. Factor 8, mastery goal orientation, measures students' goals to master learning content. Five items with factor loadings larger than 0.725 were included in this factor and Cronbach's α was 0.892. Factor 9, perceptions of computer science, addressed participants' views and self-efficacy in computer science. Seven items with factor loadings larger than 0.329 loaded on this factor and Cronbach's α was 0.768. Factor 10, STEM emotions, covered academic emotions related to STEM. Five items with factor loadings larger than 0.424 loaded on this factor and Cronbach's α was 0.702.

Reflection Card

All participants were asked to respond to a series of reflection questions in class about their (a) learning through design of robots and lessons, and (b) experience teaching the lesson. Both while learning to design robots and at the conclusion of robot design, participants were invited to reflect about the processes of their learning and design (e.g., what challenges did you have with your Ozobot programming? Explain what you did to address the challenges.). Also, participants were prompted to reflect on their teaching (e.g., What challenges did you face when using the lesson in your field experience preschool classroom? Why?). Participants were provided a total of 40 reflection questions throughout the robotics unit.

Lesson Plans Using Designed Robots

Thirty-nine participants' lesson plans were evaluated. Twenty-eight participants were teamed up as a pair and created 14 lesson plans and 11 lesson plans were created by solo work.

Procedure

Participants took the presurvey prior to the unit, and it took about 20-30 minutes. In class 1, participants were introduced to block-based robot programming and robotics in ECE. Then, they were provided a sample lesson plan, including an Ozobot Bit coding sample for preschoolers. Subsequently, they discussed the sample code and completed reflections about the sample lesson. Afterwards, participants created code in OzoBlockly by themselves and completed another reflection about the challenges in using Ozobots and programming. In field experience at local preschools after class 1, they practiced teaching the sample lesson with the given code. In class 2, reflections related to field experiences using *the sample lesson plan* were completed at the beginning. Then, participants worked on the coding task and completed reflections about their experiences with the Ozobot and OzoBlockly. After that, they were asked to use a lesson plan template to design or modify an existing lesson in one of three ways: (a) create a new lesson with new code, (b) create a new lesson with code given in class, or (c) modify the sample lesson with the codes given in the class. Subsequently, participants completed the last reflections about the process of integrating coding into their lesson plans. During field experience , participants were asked to practice teaching using the lesson plans they designed. In class 3, reflections related to field experiences using *the lesson plan they designed* were completed at the beginning. Then, participants were introduced to other kinds of robots, and they completed reflections about their further interest in robots and programming. At the end of class 3, participants took the postsurvey, which took about 20-30 minutes.

Data Analysis

Lesson Design Quality Evaluation Rubric

A modified version of *the robotic lesson plan evaluation rubric* (Authors, 2015) was used to evaluate lesson plan quality (see Section B in the supplementary material). Zero, one, or two points could be assigned to the lesson for each of 16 categories: (a) Learning goal specification, (b) Standards specification, (c) Objectives specification, (d) Key vocabulary, (e) Multi-subject inclusion, (f) STEM integration, (g) Consideration of children’s prior knowledge, (h) Listing common misconceptions, (i) Considering children’s interest, (j) Providing age-appropriate support, (k) Activity specification, (l) Providing opportunities for collaboration, (m) Closure, (n) Assessment, (o) Robot integration, (p) Programming concepts integration.

When there was no evidence in the designed lesson plan for the criterion, 0 point was given. One point (average-quality) was given to indicate that the designed lesson plan meets the criterion partially. For example, 1 point was assigned for the learning goal specification criterion when learning goals were described but not aligned with the rest of the lesson plan. Two points (high-quality) were given if the lesson plan met the criterion. For example, 2 points were given to the learning goal specification criterion when the described learning goals aligned with the designed lesson's content. The minimum possible total lesson plan score was 0, and the maximum possible score was 32. Reliability of coding of six randomly chosen lesson plans was high (intraclass correlation coefficient = 0.857).

Three raters evaluated lesson plans independently using the lesson design quality evaluation rubric and met to come to a consensus on scores. The scores for each evaluation criteria listed in the rubric were grouped into three categories – front-end analysis quality, programming and STEM integration within the lesson plan, and instructional activity quality (see Section B in the online supplement for details of the rubric sections). In the analysis, scores for each category were used as indicators of lesson plan quality.

RQ1: How can prospective teachers' lesson plan quality be classified using motivation and process variables?

Preprocessing analyses.

Sentiment analysis. The SentimentAnalysis package for R (Feuerriegel & Proellochs, 2021) was used to determine whether the students' writing was positive, negative, or neutral. The quantitative discourse analysis package (QDAP) dictionary (Rinker et al., 2014) was used in the analysis (see Section C in the supplementary material for details of sentiment analysis).

Word count. A word count was assigned to each student's responses to questions both about field experience teaching and coding challenges (Sheskin, 2011).

Analytic strategy. Multiple imputation was used to estimate missing data, which constituted less than 5% of the data. We used the MASS (Ripley et al., 2020) and KlaR (Roever et al., 2020) packages for R to conduct linear discriminant analysis (Lachenbruch & Goldstein, 1979) to predict lesson plan quality. For each predictor variable, linear discriminant coefficients were generated to form an equation to predict membership classes representing low, average, and high quality lesson plans. We created a linear combination of the following predictors that could optimally characterize the differences of three levels of lesson plan quality: : (i) Perceptions of mathematics, (ii) CS and engineering emphasis in STEM career, (iii) views of coding, (iv) Perceptions of computer science and technology, (v) Science interest, (vi) Performance goal orientation, (vii) Perceptions of English, (viii) Mastery goal orientation, (ix) Perceptions of computer science, (x) STEM emotions, (xi) Sentiment analysis – field experience teaching (see Appendix for details), (xii) Sentiment analysis – coding tasks (see Section C in the supplementary material for analytic strategy details), (xiii) Word count – field experience teaching, (xiv) Word count – coding tasks.

To understand how well motivational variables, sentiment analysis, and word count classify lesson plan quality, we investigated the classification error rate using both linear discriminant analysis and support vector machine approaches. We used a support vector machine approach with linear, radial, sigmoid, and polynomial kernels to predict lesson plan quality for each participant (Roever et al., 2020). In support vector machines, a kernel projects a set of inputs into a high dimensional space and transforms the inputs into the required form. The use of kernel can also speed up the calculation process. This resulted in a prediction of lesson plan quality, which then could be compared to the actual rated lesson plan quality. The error rate is calculated as the number of misclassified lesson plans divided by the total number of lesson plans (Efron, 1983).

RQ2: How do motivation and process variables predict prospective teachers' lesson plan quality?

We used an ordinal logistic regression approach to predict lesson plan quality using motivation and process variables. Ordinal logistic regression is used when the outcome variable is at the ordinal scale, as was the case with our lesson plan quality rating scale. Betas within logistic regression indicate the amount the outcome changes in terms of log-odds.

Results

RQ1: How can prospective teachers' lesson plan quality be classified using motivation and process variables?

Coefficients of linear discriminants for classification according to front-end analysis quality, STEM and programming integration quality, and instructional activities quality are listed in Tables 1, 2, and 3, respectively. Note that the coefficients of Linear Discriminant 1 and Linear Discriminant 2 provide the linear combination of the 14 variables that are used to form the

Table 1. Coefficients of linear discriminants for front-end analysis quality classification.

	Linear Discriminant 1	Linear Discriminant 2
Perceptions of mathematics	-0.036	-0.019
CS and engineering emphasis in STEM career	0.072	0.015
Views of coding	-0.080	-0.088
Perceptions of computer science and technology	-0.031	0.019
Science interest	-0.163	-0.031
Performance goal orientation	0.110	0.099

Perceptions of English	0.094	-0.129
Mastery goal orientation	0.225	-0.302
Perceptions of computer science	-0.007	-0.026
STEM emotions	0.291	0.126
Word count: teaching	0.185	-0.223
Word count: Ozobot	-0.173	0.361
Sentiment analysis: teaching	0.156	-3.872
Sentiment analysis: Ozobot	-2.533	-1.862

classification boundary decision. The coefficient magnitude indicates how strong a variable is in terms of determining the classification boundary for a linear discriminant (see Section D in the supplementary material for graphs demonstrating classification decision boundaries of front-end analysis quality, STEM and programming integration quality, and instructional activity quality for the first and the second linear discriminants).

The three strongest contributors for the classification decision boundary of front-end analysis quality are: sentiment score: Ozobot, STEM emotions, and mastery goal orientation for the first linear discriminant; sentiment score: Ozobot, sentiment score: teaching, and word count: Ozobot for the second linear discriminant

The three strongest contributors for the classification decision boundary of STEM and programming integration quality are: sentiment score: Ozobot, sentiment score: teaching, and

Table 2. Coefficients of linear discriminants for STEM and programming integration quality classification.

	Linear Discriminant 1	Linear Discriminant 2
Perceptions of mathematics	-0.042	-0.007
CS and engineering emphasis in STEM career	0.020	0.005
Views of coding	-0.071	-0.008
Perceptions of computer science and technology	0.039	-0.008
Science interest	0.020	0.141

Performance goal orientation	0.0301	-0.034
Perceptions of English	-0.002	-0.020
Mastery goal orientation	0.155	-0.191
Perceptions of computer science	0.024	-0.161
STEM emotions	0.240	0.019
Word count: teaching	-0.156	-0.077
Word count: Ozobot	0.217	0.049
Sentiment analysis: teaching	-3.715	11.459
Sentiment analysis: Ozobot	1.103	3.993

STEM emotions for the first linear discriminant; sentiment score: Ozobot, sentiment score: teaching, and mastery goal orientation for the second linear discriminant. The three strongest contributors for the classification decision boundary of instructional activity quality are: sentiment score: Ozobot, sentiment score: teaching, and word count: teaching for the first linear discriminant; sentiment score: Ozobot, sentiment score: teaching, and word count: Ozobot for the second linear discriminant.

We conducted classification using linear discriminant analysis and support vector machine algorithms. Note that in this section, the entire discriminant functions were used, rather than just the most important classifiers. Table 4 displays the classification error rate of linear discriminant analysis for front-end analysis quality, STEM and coding integration quality, and lowest classification error rate for front-end analysis quality, STEM and programming

Table 3. Coefficients of linear discriminants for instructional activity quality classification.

	Linear Discriminant 1	Linear Discriminant 2
Perceptions of mathematics	0.004	-0.029
CS and engineering emphasis in STEM career	0.042	0.007
Views of coding	-0.092	0.032
Perceptions of computer science and technology	0.000	-0.064
Science interest	-0.029	0.069

Performance goal orientation	0.034	-0.097
Perceptions of English	-0.023	-0.097
Mastery goal orientation	-0.123	0.286
Perceptions of computer science	-0.113	0.035
STEM emotions	0.147	-0.052
Word count: teaching	-0.185	0.188
Word count: Ozobot	0.164	-0.293
Sentiment analysis: teaching	-8.962	-2.411
Sentiment analysis: Ozobot	5.364	8.946

Table 4. Linear discriminant analysis misclassification rate.

	Front-end analysis quality	STEM and coding integration quality	Instructional activities quality
Misclassification rate	15.385%	33.333%	38.462%

integration quality, and instructional activities quality were radial, radial, and polynomial with a misclassification error rate of 12.821%, 7.692%, and 23.077%, respectively (see Section D in the supplementary material). Misclassification can be attributed to the kernel function used for the algorithm and the non-linear relationship between predictor variables and classification outcomes. We used different kernel functions to find the best input transformation functions that result in the lowest misclassification rate. The numeric variations of the kernel functions can exert an influence on classification results. Table 5 displays the classification error rate of the support vector machine with four different kernels: linear, polynomial, radial, and sigmoid. More detailed information can be found in section D of the supplementary material.

Table 5. Support vector machine misclassification rate.

	Front-end analysis quality	STEM and coding integration quality	Instructional activities quality
Linear	15.385%	25.641%	30.769%
Polynomial	28.205%	20.513%	23.077%
Radial	12.821%	7.692%	25.641%
Sigmoid	46.154%	48.718%	56.410%

Research Question #2: How do motivation and process variables predict prospective teachers' lesson plan quality?

All predictors were entered into ordinal logistic regression. For front-end analysis quality, the ordinal logistic regression deviance is 58.587 compared to the null model deviance of 84.643. For STEM and programming integration quality, the ordinal logistic regression deviance is 77.042 compared to the null model deviance of 85.074. For instructional activity quality, the ordinal logistic regression deviance is 72.753 compared to the null model deviance of 83.739.

Front-end Analysis Quality

There was one significant negative predictor for front-end analysis quality: views of coding (see Table 6). With a one-unit increase in a participant's views of coding score, we would expect a 0.202 unit decrease in the expected value of front-end analysis quality in the log-odds scale. We found two significant positive predictors: performance goal orientation and STEM emotions. With a one-unit increase in a participant's performance goal orientation score, we would expect a 0.219 unit increase in the expected value of front-end analysis quality in the log-odds scale. With a one-unit increase in a participant's STEM emotion score, we would expect a 0.371 unit increase in the expected value of front-end analysis quality in the log-odds scale.

STEM and Programming Integration Quality

The ordinal logistic regression did not detect any significant predictor variable or intercept (see Table 1 in Section D of the supplementary materials).

Table 6. Ordinal logistic model summary for front-end analysis quality.

	Beta	Standard Error	t-value	p-value
Perceptions of mathematics	-0.049	0.038	-1.267	0.205
CS and engineering emphasis in STEM career	0.072	0.040	1.794	0.073
Views of coding	-0.202	0.089	-2.273	0.023

Perceptions of computer science and technology	0.003	0.066	0.043	0.965
Science interest	-0.140	0.072	-1.936	0.053
Performance goal orientation	0.219	0.079	2.791	0.005
Perceptions of English	-0.167	0.107	-1.562	0.118
Mastery goal orientation	-0.322	0.209	-1.544	0.123
Perceptions of computer science	-0.041	0.134	-0.309	0.757
STEM emotions	0.371	0.177	2.097	0.036
Word count: teaching	-0.173	0.192	-0.900	0.368
Word count: Ozobot	0.401	0.298	1.347	0.178
Sentiment analysis: teaching	-1.057	7.962	-0.133	0.894
Sentiment analysis: Ozobot	-3.912	7.661	-0.511	0.610
Average|High	-8.688	5.602	-1.551	0.121
High|Low	-6.464	5.500	-1.175	0.240

Instructional Activity Quality

The ordinal logistic regression did not detect any significant predictor variable or intercept (see Table 2 in Section D of the supplementary materials).

Discussion

See table 7 for a summary of findings.

This research contributes to the literature in two major ways. First, it provides a vision for the use of dynamic assessment to differentiate prospective teachers who are on a path to producing high quality lesson plans and those who need further help. In this way, prospective teachers who need further support can be given such before creating lesson plans that are then used in field experience. Second, it provides insights into how learning analytics and motivation

Table 7. Summary of findings

	Findings
RQ1	a. For front-end analysis quality, sentiment score: Ozobot, STEM emotions, and mastery goal orientation were the three strongest contributors for the first linear discriminant. Sentiment score: Ozobot, sentiment score: teaching, and word

	count: Ozobot were the three strongest contributors for the second linear discriminant. b. For STEM and programming integration quality, sentiment score: Ozobot, sentiment score: teaching, and STEM emotions were the three strongest contributors for the first linear discriminant. Sentiment score: Ozobot, sentiment score: teaching, and mastery goal orientation were the three strongest contributors for the second linear discriminant. c. For instructional activity quality, sentiment score: Ozobot, sentiment score: teaching, and word count: teaching were the three strongest contributors for the first linear discriminant. Sentiment score: Ozobot, sentiment score: teaching, and word count: Ozobot were the three strongest contributors for the second linear discriminant. For front-end analysis quality, the radial kernel resulted in the lowest misclassification rate of 12.821%. For STEM and programming integration quality, the radial kernel resulted in the lowest misclassification rate of 7.692%. For instructional activities quality, the polynomial kernel resulted in the lowest misclassification rate of 23.077%.
RQ2	a. For front-end analysis quality, we found STEM emotion and performance goal orientation as significant positive predictors and views of coding as a significant negative predictor. b. For STEM and programming integration quality, we did not find any significant predictor. c. For instructional activities quality, we did not find any significant predictor.

variables may explain cognitive outcomes.

Vision for Dynamic Assessment of Prospective Teachers Learning to Plan Lessons

Creating new lesson plans is no small feat for prospective teachers (Kang et al., 2013; Lim et al., 2018; Ruys et al., 2012), even those who are already at the field experience stage (Lim et al., 2018). Within field-experience linked classes, prospective teachers are often asked to produce a lesson that they then implement in the field experience placement. This leaves little time for teacher educators to provide formative feedback before the lesson plans are used within field experience teaching. The results of the present study point to the potential for dynamic assessment of prospective teachers' work that can indicate who is on track to produce high quality lesson plans, and who needs further support. Dynamic assessment is an important strategy for supporting continual improvement, in that it can provide real time feedback on strengths and weaknesses of approaches (Kalyuga & Sweller, 2005; Swanson & Lussier, 2001).

But in this case, dynamic assessment is not meant to provide direct feedback related to the data that was assessed, but to offer additional support to those who are on track to produce low-quality lesson plans. In other words, this could lead to the provision of customized support for lesson planning based on how the equations of linear discriminants classified the quality of prospective teachers' to-be-created lesson plan. Notably, equations of linear discriminants can be used to classify prospective teachers according to lesson plan quality before they even write a lesson plan. Front-end analysis quality of lesson plans can be classified using the extent to which prospective teachers display a mastery goal orientation, the amount they write when reflecting on coding tasks, and the sentiment reflected in their reflections on coding tasks and field experience teaching (RQ1-a. in Table 7). STEM and coding integration quality of lesson plans can be classified using the amount prospective teachers write when reflecting on coding tasks, and the sentiment reflected in their reflections on coding tasks and field experience teaching (RQ1-b. in Table 7). Instructional activities quality can be classified using the amount prospective teachers write when reflecting on field experience teaching, and the sentiment reflected in their reflections on coding tasks and field experience teaching (RQ1-c. in Table 7). This can be used to identify prospective teachers who are on track and those who are not on track to integrate robotics and coding in ECE so that additional support can be provided while there is still time.

Provide Insights into how Learning Analytics and Motivation Variables may explain Cognitive Outcomes

While misclassification did occur, misclassification rates were relatively low. There was no misclassification of lesson plans predicted to be of high and low front-end analysis quality, average STEM and coding integration quality, and high and low instructional activities quality. The strongest contributors to classification of lesson plan quality were metrics that are relatively easy to gather in real time: sentiment analysis: field experience teaching, sentiment analysis: coding task, and word count: coding task. Having students write reflections in itself is a good choice pedagogically (Blomberg et al., 2013; Yost, 2006). Computers can quickly and dynamically conduct sentiment analysis and word counts on reflections on field experience teaching and completion of coding tasks, thereby indicating in real time which prospective teachers are on track to produce high quality lesson plans, and which need further support.

Of interest to motivation researchers from the goal orientation tradition are our findings regarding mastery goal orientation and performance goal orientation. Specifically, mastery goal orientation was one of the strongest contributors to classification of lessons plans' front-end analysis quality, while performance goal orientation was the only significant predictor in the ordinal logistic regression predicting front-end analysis quality (RQ2-a. in Table 7). Researchers familiar with the goal orientation literature may wonder why we refer to performance goals, rather than differentiate between performance-approach and performance avoid goals (for an overview, please see (A. Elliot & McGregor, 2001). Note that while some questions included in our survey may often load on a performance avoid goal orientation factor while others may load on a performance-avoid factor, in the factor analysis based on this dataset, they all loaded together on a common performance goal orientation factor. This is not uncommon (Brophy, 2005; Linnenbrink-Garcia et al., 2012; Urdan & Mestas, 2006), especially when participants harbor fear of failure or have low self-efficacy related to the task at hand (Linnenbrink-Garcia et al., 2012). This finding clearly needs to be unpacked. While students who hold a mastery goal approach are said to engage in learning tasks with the goal of gaining mastery in the learning content (Covington, 2000; Pintrich, 2000), much research has shown performance goals, particularly performance approach goals, to be conducive to strong learning outcomes (e.g., Harackiewicz et al., 2000; Senko et al., 2011). Furthermore, the literature shows that it is possible to hold two or more different goal orientations at the same time, and that this can be adaptive (Pintrich, 2000). While we did not find evidence of low self-efficacy or fear of failure on the part of students from the current study, we did see some evidence of dual goals of performing well and avoiding appearing incompetent.

Of the three rubric sections, front-end analysis quality is the only one for which many early childhood majors in the class would have had prior coursework. Few participants had studied (a) the integration of robotics and STEM into early childhood education, or (b) how to design a focused sequence of instructional activities to compose a lesson. But some had studied such topics as analyzing learners' prior knowledge and other characteristics. If they had exhibited a strong mastery or performance goal orientation in prior classes in which front-end analysis was covered, they likely would have learned more than other students who had neither a strong mastery orientation nor a strong performance orientation. So it makes sense that both mastery goal orientation and performance goal orientations were important to the classification or prediction of front-end analysis quality. Further research is needed to address why mastery goal orientation was a strong contributor to classification, but not prediction of, front-end analysis quality, and vice versa for performance goal orientation. Another significant classifier and positive predictor of front-end analysis quality of lesson plan was students' emotional responses to STEM (RQ2-a. in Table 7). Students who expressed positive emotions toward STEM such as excitement, enjoyment, and positive self-concept demonstrated better quality in their lesson plan front-end analysis. This result is aligned with previous literature showing a strong positive relation between academic emotions and their academic achievement (Mega et al., 2014; Murphy et al., 2019; Pekrun et al., 2002, 2017). We considered students having positive emotions towards STEM to be more likely to have higher motivation to teach STEM content and exert increased efforts to create an effective lesson plan. Particularly, they seemed to consider and specify learning goals, standards and objectives, and learner characteristics more carefully when designing robots and coding enhanced class contents and activities. Our finding also indicates that students' positive emotions toward STEM not only impact their motivation to teach better but also their actual strategies use including self-regulatory strategies and teaching strategies. This implies that considering prospective teachers' emotions helps to predict their lesson plan quality in advance and fostering positive emotions can be an effective early intervention that increases the quality of their lesson plan.

Last, we found that views of coding was a negative predictor of front-end quality of lesson plan, indicating that students' appreciation of the value of coding does not necessarily correspond to their teaching practice using coding and robots (RQ2-a. in Table 7). This finding stands in contrast to our expectations based on the expectancy-value theory that the higher perceived value of a task influence positively on performance in the task (Eccles, 2005; Wigfield & Eccles, 2000). We considered this difference in the relation between students' views of coding and quality of robot and coding enhanced lesson plan to be partially related to their STEM teacher professional identities. Considering they are undergraduate students who are learning to teach STEM education, their STEM teacher professional identities are not fully formed yet. They're more likely undergoing the process of developing their own professional identities, with having multiple identities such as leaners, curriculum designers, or collaborators.

When it comes to planning and practicing their STEM teaching, many students identified themselves as more of an "interested but confused learner" (Jiang et al., 2021, p. 9). That is, even though students understand the values of coding knowledge and skills and they believe it is worth pursuing following their career, when it comes for them to be actually asked to integrate them into their teaching, they still feel less confident as effective teachers teaching with coding. We found that many students, even when they perceived high task value, experienced challenges and difficulties over the course of creating lesson plan using coding and robots. Particularly, they felt confused and unsure about what they were doing while creating lesson plans when they found mismatches between what they think STEM education would be and how they implement their stem education ideas based on traditional learning standards, objectives, and learners' prior knowledge. This possibly caused them to think they are not ready to meet the expectations or the requirements regarding creating STEM incorporated lesson plan. In turn, this could negatively affect their performance in creating a quality lesson plan. As such, understanding prospective teachers' affective aspects and explicitly addressing them could be a critical strategy to improve their lesson plan quality, especially considering that anxiety regarding STEM and coding is more prevalent in female students and the majority of prospective teachers are female (Pelch, 2018).

Implications

Implications for Women in Computer Science

Our results point to the potential for dynamic assessment of prospective teachers' work and identified factors that are relevant to their work quality. These together can support continual improvement in prospective teachers' integrating robotics and coding in ECE by helping them design a high-quality computer science-enhanced lesson plan. Our participants were exclusively women, which is in line with preservice ECE education as a whole and the ECE field in general (Laere et al., 2014). This means that some of our findings have potential implications for women in the computer science field. Notably, women are severely underrepresented in computer science: women account for only around 20% of undergraduates graduating with a degree in computer science, and around 25 % of those employed in computer related occupations (Beyer, 2014). This has been partially attributed to stereotype threat, in which women often perform worse on tests of computer science than their skills would predict because they are reminded of stereotypes that computer science is for men (Cheryan & Markus, 2020; Thoman et al., 2013).

Previous studies also report that women often have low computer self-efficacy, which in turn lowers female students' interest in the discipline and motivation to do well (Beyer, 2014). They are less likely to identify with computer science and have a low sense of belongings and expectancy for success in computer science (Falkner et al., 2015). By helping early childhood prospective teachers to better design robot and coding enhanced lesson plans, female prospective teachers can have increased confidence in teaching computer science concepts and skills, which will increase a chance for them to incorporate more CS content and activities into their teaching. Furthermore, this will help increase the number of adequate role models for female students by providing examples of women who have high confidence and efficacy in dealing with robots and coding. This can impact tremendously the students at an early age and can instill confidence and interest particularly in girls who are learning computer science. Much literature has shown that one of the most critical reasons for women's low interest and self-efficacy in CS is the lack of role models (Cheryan et al., 2011; Vitores & Gil-Juárez, 2016).

Implications for Teacher Educators

Dynamic assessment of prospective teachers learning to plan lessons requires that teacher educators be ready and willing to provide additional support to those prospective teachers who are deemed to be on track to produce a low-quality lesson plan. Indicating which students need additional support may be done through the use of a teacher educator dashboard (Park & Jo, 2015). This will allow teacher educators to provide timely and individualized feedback to support prospective teachers. This can help ECE prospective teachers focus on the points that need more attention (Korth & Baum, 2011). As the system is used more extensively, the accuracy of the model will increase, thereby increasing the utility of the system.

Implications for Prospective Teachers as Learners and Designers

While the discriminant functions can indicate which prospective teachers are on track to produce a low-quality lesson plan, and thus are in need of additional support, it can also indicate to prospective teachers areas where additional effort may be required. That is, with timely individualized feedback on their reflections, prospective teachers can be informed of what could be the weakest design point that needs improvement in the lesson plan. Prospective teachers can be timely notified of the reflections' content, word counts, and sentiment scores in the process of learning and design, which can enable them to mindfully monitor their lesson plan writing process. This feedback may in turn generate the dispositions to engage in deeper reflection on the part of prospective teachers, which then may be carried into their futures as practicing teachers. This in turn would maximize their potential as teachers as designers.

Limitations and Suggestions for Future Research

Power was limited due to the relatively small sample size. This caused misclassification rates to be higher than hoped, and for the logistic regression equations to have a relatively small number of significant predictors. Still, the fact that the vast majority of lesson plans were classified correctly in terms of front-end analysis quality, STEM and coding integration quality, and instructional activities quality is noteworthy and provides good direction for future research.

There are several aspects that limit the generalizability of our findings. First of all, all participants of our study were female and most of them were white. Because of the distinct features of the early-childhood prospective teacher population (i.e., disproportionately female and white population), our findings are likely to be generalized to those populations. However, certainly our results are confined to white female prospective teacher populations and may not be generalized to male and other ethnicity/race groups (e.g., Latinx and African American). In addition, our data were collected from a single institution. Considering that there exist differences in institutional characteristics including student characteristics, our data limits our ability to generalize findings to the total population of prospective teachers. As such, future study needs to include prospective teachers who are from more diverse educational and socioeconomic backgrounds. Also, a broader exploration of multiple variables that are relevant to prospective teachers who are male and underrepresented minorities in the STEM+C field is much needed.

Conclusion

Learning to plan high-quality lessons is difficult for prospective teachers. In the present study we discovered how front-end analysis quality, STEM and coding integration quality, and instructional activities quality of lesson plans could be classified and predicted. Misclassification rates were relatively low. These results have important implications for dynamic assessment of prospective teachers engaging in lesson planning, indicating which prospective teachers are on track to produce high quality lessons, and which ones need further support.

Acknowledgements

This research was supported by grants 1906059 and 1927595 from the National Science Foundation (NSF). Any opinions, findings, or conclusions are those of the authors and do not necessarily represent official viewpoints of NSF. This research was designed by the authors and the funder had no influence on the design, analysis, or interpretation.

References

Akkerman, S. F., & Meijer, P. C. (2011). A dialogical approach to conceptualizing teacher identity. *Teaching and Teacher Education*, *27*(2), 308–319. https://doi.org/10.1016/j.tate.2010.08.013

Aras, S. (2016). Free play in early childhood education: A phenomenological study. *Early Child Development and Care*, *186*(7), 1173–1184. https://doi.org/10.1080/03004430.2015.1083558

Arndt, S. (2018). Early childhood teacher cultural Otherness and belonging. *Contemporary Issues in Early Childhood*, *19*(4), 392–403. https://doi.org/10.1177/1463949118783382

Asensio-Pérez, J. I., Dimitriadis, Y., Pozzi, F., Hernández-Leo, D., Prieto, L. P., Persico, D., & Villagrá-Sobrino, S. L. (2017). Towards teaching as design: Exploring the interplay between full-lifecycle learning design tooling and Teacher Professional Development. *Computers & Education*, *114*, 92–116. https://doi.org/10.1016/j.compedu.2017.06.011

Ashiabi, G. S. (2007). Play in the preschool classroom: Its socioemotional significance and the teacher's role in play. *Early Childhood Education Journal*, *35*(2), 199–207. https://doi.org/10.1007/s10643-007-0165-8

Backfisch, I., Lachner, A., Hische, C., Loose, F., & Scheiter, K. (2020). Professional knowledge or motivation? Investigating the role of teachers' expertise on the quality of technology-enhanced lesson plans. *Learning and Instruction*, *66*, 101300. https://doi.org/10.1016/j.learninstruc.2019.101300

Beijaard, D. (2017). Learning teacher identify in teacher education. In D. J. Clandinin & J. Husu (Eds.), *The SAGE Handbook of Research on Teacher Education*. SAGE Publications. http://ebookcentral.proquest.com/lib/pensu/detail.action?docID=4880685

Bers, M. U. (2008). *Blocks to robots: Learning with technology in the early childhood classroom*. Teachers College Press.

Beyer, S. (2014). Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education*, *24*. https://doi.org/10.1080/08993408.2014.963363

Blomberg, G., Sherin, M. G., Renkl, A., Glogger, I., & Seidel, T. (2013). Understanding video as a tool for teacher education: Investigating instructional strategies to promote reflection. *Instructional Science*, *42*(3), 443–463. https://doi.org/10.1007/s11251-013-9281-6

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*, 61–100. https://doi.org/10.2307/1167267

Brooker, E., Blaise, M., & Edwards, S. (Eds.). (2014). *SAGE Handbook of Play and Learning in Early Childhood*. SAGE.

Brophy, J. (2005). Goal theorists should move on from performance goals. *Educational Psychologist*, *40*(3), 167–176. https://doi.org/10.1207/s15326985ep4003_3

Çetin, M., & Demircan, H. Ö. (2020). Empowering technology and engineering for STEM education through programming robots: A systematic literature review. *Early Child*

Development and Care, *190*(9), 1323–1335.
https://doi.org/10.1080/03004430.2018.1534844

Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review*, *127*(6), 1022–1052.
http://dx.doi.org.ezaccess.libraries.psu.edu/10.1037/rev0000209

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do Female and Male Role Models Who Embody STEM Stereotypes Hinder Women's Anticipated Success in STEM? *Social Psychological and Personality Science*, *2*(6), 656–664.
https://doi.org/10.1177/1948550611405218

Cober, R., Tan, E., Slotta, J., So, H.-J., & Könings, K. D. (2015). Teachers as participatory designers: Two case studies with technology-enhanced learning environments. *Instructional Science: An International Journal of the Learning Sciences*, *43*(2), 203–228.

Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology*, *51*(1), 171–200.
https://doi.org/10.1146/annurev.psych.51.1.171

Davis, E. A. (2006). Characterizing productive reflection among preservice elementary teachers: Seeing what matters. *Teaching and Teacher Education*, *22*(3), 281–301.
https://doi.org/10.1016/j.tate.2005.11.005

Di Lieto, M. C., Inguaggiato, E., Castro, E., Cecchi, F., Cioni, G., Dell'Omo, M., Laschi, C., Pecini, C., Santerini, G., Sgandurra, G., & Dario, P. (2017). Educational robotics intervention on executive functions in preschool children: A pilot study. *Computers in Human Behavior*, *71*, 16–23. https://doi.org/10.1016/j.chb.2017.01.018

Ding, M., & Carlson, M. A. (2013). A use of the IES recommendations: A use of the ies recommendations. *The Elementary School Journal*, *113*(3), 359–385. JSTOR. https://doi.org/10.1086/668505

Dolianiti, F. S., Iakovakis, D., Dias, S. B., Hadjileontiadou, S., Diniz, J. A., & Hadjileontiadis, L. (2019). Sentiment analysis techniques and applications in education: A survey. In M. Tsitouridou, J. A. Diniz, & T. A. Mikropoulos (Eds.), *Technology and Innovation in Learning, Teaching and Education* (pp. 412–427). Springer International Publishing. https://doi.org/10.1007/978-3-030-20954-4_31

Doppelt, Y. (2009). Assessing creative thinking in design-based learning. *International Journal of Technology and Design Education*, *19*(1), 55–65. https://doi.org/10.1007/s10798-006-9008-y

Eccles, J. S. (2005). Subjective task value and the Eccles et al. Model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of Competence and Motivation, First Edition* (pp. 105–121). Guilford Publications.

Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association*, *78*(382), 316–331. https://doi.org/10.2307/2288636

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

Falkner, K., Szabo, C., Michell, D., Szorenyi, A., & Thyer, S. (2015). Gender Gap in Academia: Perceptions of Female Computer Science Academics. *Proceedings of the 2015 ACM*

Conference on Innovation and Technology in Computer Science Education, 111–116. https://doi.org/10.1145/2729094.2742595

Feuerriegel, S., & Proellochs, N. (2021). *SentimentAnalysis R package*. https://github.com/sfeuerriegel/SentimentAnalysis

Foley, L. S. (2011). Exploring K–3 Teachers' Implementation of Comprehension Strategy Instruction (CSI) Using Expectancy-Value Theory. *Literacy Research and Instruction*, *50*(3), 195–215. https://doi.org/10.1080/19388071.2010.505680

Friesen, M. D., & Besley, S. C. (2013). Teacher identity development in the first year of teacher education: A developmental and social psychological perspective. *Teaching and Teacher Education*, *36*, 23–32. https://doi.org/10.1016/j.tate.2013.06.005

Fulmer, S. M., & Frijters, J. C. (2009). A review of self-report and alternative approaches in the measurement of student motivation. *Educational Psychology Review*, *21*(3), 219–246. https://doi.org/10.1007/s10648-009-9107-x

Grover, S., & Basu, S. (2017). Measuring student learning in introductory block-based programming: Examining misconceptions of loops, variables, and boolean logic. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 267–272. https://doi.org/10.1145/3017680.3017723

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology*, *92*(2), 316–330. https://doi.org/10.1037/0022-0663.92.2.316

Hofer, B. K. (2017). Shaping the epistemology of teacher practice through reflection and reflexivity. *Educational Psychologist*, *52*(4), 299–306. https://doi.org/10.1080/00461520.2017.1355247

Hunsaker, E. (2018). Ozobot Bit: A Guide for Parents & Educators. *Brigham Young University*.

Jaipal-Jamani, K., & Angeli, C. (2018). Developing teacher self-efficacy to teach science and computational thinking with educational robotics: Using scaffolded programming scripts. In C. B. Hodges (Ed.), *Self-Efficacy in Instructional Technology Contexts* (pp. 183–203). Springer International Publishing. https://doi.org/10.1007/978-3-319-99858-9_11

Janssen, N., & Lazonder, A. W. (2016). Supporting pre-service teachers in designing technology-infused lesson plans. *Journal of Computer Assisted Learning*, *32*(5), 456–467. https://doi.org/10.1111/jcal.12146

Jiang, H., Wang, K., Wang, X., Lei, X., & Huang, Z. (2021). Understanding a STEM teacher's emotions and professional identities: A three-year longitudinal case study. *International Journal of STEM Education*, *8*(1), 51. https://doi.org/10.1186/s40594-021-00309-9

John, P. D. (2006). Lesson planning and the student teacher: Re-thinking the dominant model. *Journal of Curriculum Studies*, *38*(4), 483–498. https://doi.org/10.1080/00220270500363620

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36.

Kali, Y., McKenney, S., & Sagy, O. (2015). Teachers as designers of technology enhanced learning. *Instructional Science*, *43*(2), 173–179. https://doi.org/10.1007/s11251-014-9343-4

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development*, *53*(3), 83–93. https://doi.org/10.1007/BF02504800

Kang, E. J. S., Bianchini, J. A., & Kelly, G. J. (2013). Crossing the border from science student to science teacher: Preservice teachers' views and experiences learning to teach inquiry. *Journal of Science Teacher Education*, *24*(3), 427–447. https://doi.org/10.1007/s10972-012-9317-9

Kidd, J., Kaipa, K., Sacks, S., & Almeida, L. M. de S. (2020). *Introducing coding into teacher education: An interdisciplinary robotics experience for education and engineering students*. 1303–1310. https://www.learntechlib.org/primary/p/216040/

Kim, C., Kim, D., Yuan, J., Hill, R. B., Doshi, P., & Thai, C. N. (2015). Robotics to promote elementary education pre-service teachers' STEM engagement, learning, and teaching. *Computers & Education*, *91*, 14–31. https://doi.org/10.1016/j.compedu.2015.08.005

Koehler, M. J., Mishra, P., & Yahya, K. (2007). Tracing the development of teacher knowledge in a design seminar: Integrating content, pedagogy and technology. *Computers & Education*, *49*(3), 740–762.

König, J., Bremerich-Vos, A., Buchholtz, C., & Glutsch, N. (2020). General pedagogical knowledge, pedagogical adaptivity in written lesson plans, and instructional practice among preservice teachers. *Journal of Curriculum Studies*, *0*(0), 1–23. https://doi.org/10.1080/00220272.2020.1752804

Kopcha, T. J. (2007). Self-presentation bias in surveys of teacher's educational technology practices. *Educational Technology Research and Development*, *55*(6), 627–646.

Korth, B. B., & Baum, A. C. (2011). Teachers supporting future teachers: A critical part of early childhood teacher preparation. *YC Young Children*, *66*(3), 20.

Kwo, O. (Ed.). (2010). *Teachers as learners: Critical discourse on challenges and opportunities*. Springer Netherlands. https://doi.org/10.1007/978-90-481-9676-0

Lachenbruch, P. A., & Goldstein, M. (1979). Discriminant analysis. *Biometrics*, 69–85. https://doi.org/10.2307/2529937

Laere, K. V., Vandenbroeck, M., Roets, G., & Peeters, J. (2014). Challenging the feminisation of the workforce: Rethinking the mind–body dualism in Early Childhood Education and Care. *Gender and Education*, *26*(3), 232–245. https://doi.org/10.1080/09540253.2014.901721

Laurillard, D. (2012). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.

Lee, C.-J., & Kim, C. (2014). An implementation study of a TPACK-based instructional design model in a technology integration course. *Educational Technology Research and Development*, *62*(4), 437–460. https://doi.org/10.1007/s11423-014-9335-8

Lee, Y., & Lee, J. (2014). Enhancing pre-service teachers' self-efficacy beliefs for technology integration through lesson planning practice. *Computers & Education*, *73*, 121–128. https://doi.org/10.1016/j.compedu.2014.01.001

Lee, Y.-A., & Takahashi, A. (2011). Lesson plans and the contingency of classroom interactions. *Human Studies*, *34*(2), 209–227. https://doi.org/10.1007/s10746-011-9181-1

Liang, J.-C., & Tsai, C.-C. (2008). Internet self-efficacy and preferences toward constructivist internet-based learning environments: A study of pre-school teachers in Taiwan. *Educational Technology & Society*, *11*(1), 226–237.

Lim, W., Son, J.-W., & Kim, D.-J. (2018). Understanding preservice teacher skills to construct lesson plans. *International Journal of Science and Mathematics Education*, *16*(3), 519–538. https://doi.org/10.1007/s10763-016-9783-1

Linnenbrink-Garcia, L., Middleton, M., Ciani, K., Easter, M., O'Keefe, P., & Zusho, A. (2012). The strength of the relation between performance-approach and performance-avoidance goal orientations: Theoretical, methodological, and instructional implications. *Educational Psychologist*, *47*, 281–301. https://doi.org/10.1080/00461520.2012.722515

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, *5*(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Martinez, A., Agirre, N., López-de-Arana, E., & Bilbatua, M. (2019). Analysis of interaction patterns and tutor assistance in processes of joint reflection in pre-service teacher education. *Journal of Education for Teaching*, *45*(4), 389–401. https://doi.org/10.1080/02607476.2019.1639259

Matuk, C. F., Linn, M. C., & Eylon, B.-S. (2015). Technology to support teachers using evidence from student work to customize technology-enhanced inquiry units. *Instructional Science: An International Journal of the Learning Sciences*, *43*(2), 229–257.

McKenney, S., Kali, Y., Markauskaite, L., & Voogt, J. (2015). Teacher design knowledge for technology enhanced learning: An ecological framework for investigating assets and needs. *Instructional Science*, *43*(2), 181–202. https://doi.org/10.1007/s11251-014-9337-2

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology*, *106*(1), 121–131. https://doi.org/10.1037/a0033546

Middlewood, D., Parker, R., & Beere, J. (2005). *Creating a learning school*. SAGE Publications. http://ebookcentral.proquest.com/lib/pensu/detail.action?docID=343985

Milner, A. R., Templin, M. A., & Czerniak, C. M. (2011). Elementary science students' motivation and learning strategy use: Constructivist classroom contextual factors in a life science laboratory and a traditional classroom. *Journal of Science Teacher Education*, *22*(2), 151–170. http://dx.doi.org.ezaccess.libraries.psu.edu/10.1007/s10972-010-9200-5

Mintzes, J. J., Marcum, B., Messerschmidt-Yates, C., & Mark, A. (2013). Enhancing self-efficacy in elementary science teaching with professional learning communities. *Journal of Science Teacher Education*, *24*(7), 1201–1218. https://doi.org/10.1007/s10972-012-9320-1

Molee, L. M., Henry, M. E., Sessa, V. I., & Mckinney-Prupis, E. R. (2011). Assessing Learning in Service-Learning Courses through Critical Reflection. *Journal of Experiential Education*, *33*(3), 239–257. https://doi.org/10.1177/105382590113300304

Murphy, S., MacDonald, A., Wang, C. A., & Danaia, L. (2019). Towards an Understanding of STEM Engagement: A Review of the Literature on Motivation and Academic Emotions. *Canadian Journal of Science, Mathematics and Technology Education*, *19*(3), 304–320. https://doi.org/10.1007/s42330-019-00054-w

Musselwhite, D. J., & Wesolowski, B. C. (2018). Evaluating the psychometric qualities of a rating scale to assess pre-service teachers' lesson plan development in the context of a secondary-level music performance classroom. *Journal of Research in Music Education*, *66*(3), 338–358. https://doi.org/10.1177/0022429418793645

Näykki, P., Kontturi, H., Seppänen, V., Impiö, N., & Järvelä, S. (2021a). Teachers as learners – a qualitative exploration of pre-service and in-service teachers' continuous learning community OpenDigi. *Journal of Education for Teaching*, *47*(4), 495–512. https://doi.org/10.1080/02607476.2021.1904777

Näykki, P., Kontturi, H., Seppänen, V., Impiö, N., & Järvelä, S. (2021b). Teachers as learners: A qualitative exploration of pre-service and in-service teachers' continuous learning community OpenDigi. *Journal of Education for Teaching*, *47*(4), 495–512. https://doi.org/10.1080/02607476.2021.1904777

Ni, L., & Guzdial, M. (2012). Who AM I? Understanding high school computer science teachers' professional identity. *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 499–504. https://doi.org/10.1145/2157136.2157283

Olsen, B. (2008). Introducing teacher identity and this volume. *Teacher Education Quarterly*, *35*(3), 3–6.

Ozogul, G., Olina, Z., & Sullivan, H. (2008). Teacher, self and peer evaluation of lesson plans written by preservice teachers. *Educational Technology Research and Development*, *56*(2), 181.

Park, Y., & Jo, I.-H. (2015). Development of the Learning Analytics Dashboard to Support Students' Learning Performance. *Journal of Universal Computer Science*, *21*(1), 110–133.

Parker, P. D., Martin, A. J., Colmar, S., & Liem, G. A. (2012). Teachers' workplace well-being: Exploring a process model of goal orientation, coping behavior, engagement, and burnout. *Teaching and Teacher Education*, *28*(4), 503–513. https://doi.org/10.1016/j.tate.2012.01.001

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. In *National Center for Education Research*. National Center for Education Research. https://eric.ed.gov/?id=ED498555

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. *Educational Psychologist*, *37*(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement Emotions and Academic Performance: Longitudinal Models of Reciprocal Effects. *Child Development*, *88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Pelch, M. (2018). Gendered differences in academic emotions and their implications for student success in STEM. *International Journal of STEM Education*, *5*(1), 33. https://doi.org/10.1186/s40594-018-0130-7

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, *92*, 544–555. https://doi.org/10.1037/0022-0663.92.3.544

Plate, E., & Peacock, A. (2021). Teachers' learning in schools moving towards inclusion: Experiences from england and germany. In A. Köpfer, J. J. W. Powell, & R. Zahnd (Eds.), *Handbuch Inklusion international / International Handbook of Inclusive Education* (1st ed., pp. 423–452). Verlag Barbara Budrich. https://doi.org/10.2307/j.ctv1f70kvj.26

Pringle, R. M., Dawson, K., & Ritzhaupt, A. D. (2015). Integrating science and technology: Using technological pedagogical content knowledge as a framework to study the practices of science teachers. *Journal of Science Education and Technology*, *24*(5), 648–662. https://doi.org/10.1007/s10956-015-9553-9

Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? *Review of Educational Research*, *82*(3), 330–348. https://doi.org/10.3102/0034654312457429

Rinker, T., Gillespie, C., & Citro, C. (2014). *Qdap Version 2.1.0*. Zenodo. https://doi.org/10.5281/ZENODO.11124

Ripley, B., Venables, B., Bates, D. M., ca 1998), K. H. (partial port, ca 1998), A. G. (partial port, & Firth, D. (2020). *MASS: Support Functions and Datasets for Venables and Ripley's MASS* (7.3-51.6) [Computer software]. https://CRAN.R-project.org/package=MASS

Roever, C., Raabe, N., Luebke, K., Ligges, U., Szepannek, G., Zentgraf, M., & Meyer, D. (2020). *klaR: Classification and Visualization* (0.6-15) [Computer software]. https://CRAN.R-project.org/package=klaR

Røise, Ø., Edeholt, H., Morrison, A., Bjørkli, C., & Hoff, T. (2014). What we talk about when we talk about design. *FORMakademisk*, *7*(2), 1–17.

Runhaar, P., Sanders, K., & Yang, H. (2010). Stimulating teachers' reflection and feedback asking: An interplay of self-efficacy, learning goal orientation, and transformational leadership. *Teaching and Teacher Education*, *26*(5), 1154–1161. https://doi.org/10.1016/j.tate.2010.02.011

Ruys, I., Keer, H. V., & Aelterman, A. (2012). Examining pre-service teacher competence in lesson planning pertaining to collaborative learning. *Journal of Curriculum Studies*, *44*(3), 349–379. https://doi.org/10.1080/00220272.2012.675355

Sak, R. (2015). Comparison of self-efficacy between male and female pre-service early childhood teachers. *Early Child Development and Care*, *185*(10), 1629–1640. https://doi.org/10.1080/03004430.2015.1014353

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*, 26–47. https://doi.org/10.1080/00461520.2011.538646

Sheskin, D. J. (2011). *Handbook of parametric and nonparametric statistical procedures* (5th ed.). CRC Press.

Shulman, L. S., & Sherin, M. G. (2004). Fostering communities of teachers as learners: Disciplinary perspectives. *Journal of Curriculum Studies*, *36*(2), 135–140. https://doi.org/10.1080/0022027032000135049

Sias, C. M., Nadelson, L. S., Juth, S. M., & Seifert, A. L. (2017). The best laid plans: Educational innovation in elementary teacher generated integrated STEM lesson plans. *The Journal of Educational Research*, *110*(3), 227–238. https://doi.org/10.1080/00220671.2016.1253539

Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports*, *114*(1), 68–77. https://doi.org/10.2466/14.02.PR0.114k14w0

Sullivan, A., & Bers, M. U. (2019). Investigating the use of robotics to increase girls' interest in engineering during early elementary school. *International Journal of Technology and Design Education*, *29*(5), 1033–1051. https://doi.org/10.1007/s10798-018-9483-y

Svihla, V., Reeve, R., Sagy, O., & Kali, Y. (2015). A fingerprint pattern of supports for teachers' designing of technology-enhanced learning. *Instructional Science: An International Journal of the Learning Sciences*, *43*(2), 283–307.

Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*, *71*, 321–363. https://doi.org/10.3102/00346543071002321

Thoman, D. B., Smith, J. L., Brown, E. R., Chase, J., & Lee, J. Y. K. (2013). Beyond performance: A motivational experiences model of stereotype threat. *Educational Psychology Review*, *25*(2), 211–243. https://doi.org/10.1007/s10648-013-9219-1

Toh, L. P. E., Causo, A., Tzuo, P.-W., Chen, I.-M., & Yeo, S. H. (2016). A review on the use of robots in education and young children. *Journal of Educational Technology & Society*, *19*(2), 148–163.

Torres, N., González, R., & Carvalho, J. L. (2018). Development of basic spatial notions through work with educational robotics in the early childhood education classroom and analysis of qualitative data with webqda software. In A. Costa, L. Reis, & A. Moreira (Eds.), *Computer Supported Qualitative Research: New Trends on Qualitative Research* (Vol. 861, p. 22). Springer.

Tsai, C.-C., & Chai, C. S. (2012). The "third"-order barrier for technology-integration instruction: Implications for teacher education. *Australasian Journal of Educational Technology*, *28*(6), 1057–1060.

Tsai, P.-S., & Tsai, C.-C. (2019). Preservice teachers' conceptions of teaching using mobile devices and the quality of technology integration in lesson plans. *British Journal of Educational Technology*, *50*(2), 614–625. https://doi.org/10.1111/bjet.12613

Urdan, T., & Mestas, M. (2006). The goals behind performance goals. *Journal of Educational Psychology*, *98*(2), 354–365. https://doi.org/10.1037/0022-0663.98.2.354

Vitores, A., & Gil-Juárez, A. (2016). The trouble with 'women in computing': A critical examination of the deployment of research on the gender gap in computer science. *Journal of Gender Studies*, *25*(6), 666–680. https://doi.org/10.1080/09589236.2015.1087309

Voogt, J., Laferrière, T., Breuleux, A., Itow, R. C., Hickey, D. T., & McKenney, S. (2015). Collaborative design as a form of professional development. *Instructional Science: An International Journal of the Learning Sciences*, *43*(2), 259–282.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wen, M., Yang, D., & Rosé, C. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us? *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)*, 130–137.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.

Wulff, P., Buschhüter, D., Westphal, A., Nowak, A., Becker, L., Robalino, H., Stede, M., & Borowski, A. (2021). Computer-based classification of preservice physics teachers' written reflections. *Journal of Science Education and Technology*, *30*(1), 1–15. https://doi.org/10.1007/s10956-020-09865-1

Yost, D. S. (2006). Reflection and self-efficacy: Enhancing the retention of qualified teachers from a teacher education perspective. *Teacher Education Quarterly*, *33*(4), 59–76.

Zviel-Girshin, R., Luria, A., & Shaham, C. (2020). Robotics as a tool to enhance technological thinking in early childhood. *Journal of Science Education and Technology*, *29*(2), 294–302. https://doi.org/10.1007/s10956-020-09815-x