Michael Morrone, Editor-in-Chief
Christopher J. Young, Managing Editor
Steve Rahko, Editor
Karissa Rector, Editorial Assistant

JoSoTL
Journal of the Scholarship of Teaching and Learning

Social Justice in Action: A Document Analysis of the Integration of Social Justice Principles into Teaching

Emily Spitzman
Bridgewater State University
espitzman@bridgew.edu

Alexandra Balconi
Bridgewater State University

Abstract: Teachers of all students, particularly English Learners (ELs), need to integrate social justice pedagogy into their lessons so that all learners are included in the learning process, thinking critically about curriculum and taking action in the face of injustice. There has been some research into teacher preparation programs focusing on how they integrate culturally responsive and social justice pedagogy into their curricula and whether there has been a positive impact on teachers' self-awareness, social justice knowledge and classroom practices as a result (Ruffin, 2016; Thieman, 2016). However, these studies do not address lesson content. This document analysis study, framed theoretically with critical intercultural communication (Halualani & Nakayama, 2010), explored the integration of social justice principles into lesson plans developed by pre-service and in-service English as a second language (ESL) teachers who were pursuing a TESOL graduate degree (Initial License) at a university in the Northeast of the United States. The lesson plans were analyzed using a rubric aligned with the Social Justice Standards: The Teaching Tolerance Anti-Biases Framework. The exploration unveiled the need for more connections to students' backgrounds, structured in-class dialogues, support for linguistic needs and modeling of intercultural practices.

Keywords: social justice standards, dialogue, intercultural practices, bias

The concept of social justice has become increasingly common in education. More and more educational institutions, programs and organizations integrate social justice orientations into their missions and practices. Social justice in education is characterized by respectful communication and inclusive approaches to teaching and learning that can lead to more equitable educational outcomes (Nieto & Bode, 2018). However, how exactly can social justice be integrated into such domains of teaching as lesson planning, delivering instruction and assessing student learning? What does it entail and how are we preparing pre-service and in-service teachers in academic degree programs to design lesson plans and deliver instruction for social justice? Do they know what a lesson needs to include in order to be considered a lesson grounded in social justice principles? What are these principles? And how do they inform teaching? All of these questions need to be addressed in teacher education preparation. It is important to discuss and promote social justice principles and their application across all education tracks and ensure that it is not limited to multicultural education programs and courses. Teaching English to Speakers of Other Languages (TESOL) academic programs are not an exception considering the fact that diversity is a crucial characteristic of a classroom of English learners (ELs).

It is not sufficient to follow a prescribed curriculum; instead it is necessary to plan instruction to address social realities and dynamics within the teaching context and beyond (Ciechanowski, 2013). This entails integrating practices focusing on inclusion, respectful relationships, community building

and action against bias and injustice. These social justice principles must be a solid component of instructional practices, activities, and routines.

The purpose of the document analysis described in this article was to examine lessons created by pre-service and in-service teachers in a TESOL graduate program for initial K-12 licensure at a university in New England. Particularly, the integration of social justice principles into those lesson plans was examined. To define social justice principles, *Social Justice Standards: The Teaching Tolerance Anti-Biases Framework* (Teaching Tolerance, 2016) and related literature on social justice in education described in depth below were used for the analysis.

Literature Review

This literature review begins with an overview of the theoretical framework and guiding philosophy for this study, critical intercultural communication (Halualani & Nakayama, 2010) and the multilingual turn (May, 2014), and then discusses definitions of social justice and methods of integrating social justice into teaching.

Theoretical Framework and Guiding Philosophy

This study is framed theoretically with critical intercultural communication (Halualani & Nakayama, 2010), a framework that purports that all communication is power-laden; it critiques the notion that there is a level playing field across cultures. Intercultural communication studies have often ignored power imbalances in communication, even though this is almost, arguably always, impossible to do. When two people communicate across differences, there are power dynamics at play and those dynamics must be integrated into the way that communication is examined and understood. Piller (2016) discusses the idea that linguistic diversity is a term that "has in some contexts become a euphemism for linguistic subordination" (p. 7). Dominant and subordinate identities are interacting within all individuals depending on the context, and it is essential that the study of these power-laden dynamics is integrated in teacher education programs, as is explained below.

In critical applied linguistics, these power dynamics play a crucial role, and there has been a turn away from the native/non-native dichotomy where language learners are seen as lacking. Instead, the discipline underscores the competencies of bi/multilingual learners (May, 2014). The fields of TESOL and SLA (Second Language Acquisition) have been critiqued for lagging behind this multilingual turn and yet there are a variety of voices trying to change the dominant narratives. Canagarajah (2014) explains the multilingual repertoire that learners develop as they are using languages. Instead of viewing languages as distinct, this is an integrated view of language learning that highlights the strengths and capacities of multilingual learners. Norton (2014) describes the role of investment in language learning, suggesting that one's social, cultural and linguistic identities must be integrated in order for learners to fully commit to the learning process.

Social Justice in Teaching and Learning

Teaching and learning that account for the power imbalances inherent in communication and the multi-faceted identities of individuals are critical so that *all* students are able to learn from the curriculum and from one another. The current study uses Nieto and Bode's (2018) definition of social justice: "a philosophy, an approach, and actions that embody treating all people with fairness, respect, dignity, and generosity. On a societal scale, this means affording each person the real - not simply a stated or codified - opportunity to achieve to her or his potential and full participation in a democratic society by giving each person access to the goods, services and cultural capital of a society, while also

affirming the culture and talents of each individual and the group or groups with which she or he identifies" (Nieto & Bode, 2018, p. 8). A classroom community should not only foster fair and generous communication among all students, but it also needs to ensure that learners are integrated into interactions, with equitable access to curricula. This is particularly critical in a classroom of English language learners as a lack of language proficiency may restrict them from classroom participation if proper supports are not provided.

Education for social justice is not a new idea and many are working tirelessly to ensure that education is built on equity, activism and social literacy (Ayers, Quinn, & Stoval, 2009, p. xiv). Education has historically been deeply inequitable in the United States, with severe racial and economic disparities in educational quality and outcomes. The curriculum offered to many children in U.S. schools, especially students of color and those from lower socioeconomic backgrounds, has focused on lower-order "rote" skills, not asking children to engage with one another and problem solve. To counteract this and include all students in critical thinking, it is essential that teachers are prepared to teach higher-order thinking skills to all students (Darling-Hammond, 2010).

Integration of Social Justice into Classroom Communities

In effective education for social justice, all students are integrated into the classroom, learning from the curriculum. There are myriad opportunities for critical thinking and self-reflection, questioning power imbalances and the status quo as well as taking action in the face of injustice. This section presents an exploration of the research on the components of education for social justice, pointing to research gaps and the need for the current study.

In order for all students to reach their academic potential, reciprocal teaching approaches, where power is shared by teachers and students, are essential. Establishing connections to students' backgrounds, experiences and prior learning is consistent with social justice principles and plays a critical role in students' learning (Ciechanowski, 2013). These connections make learning meaningful and authentic, thus increasing opportunities for understanding the concepts taught in the lesson and providing a solid foundation for future learning (Echevarría & Graves, 2010). Teachers should include students' stories in lessons when possible, as life stories foster engagement in instruction and become the building blocks of effective planning and curriculum. They expose both ELs and non-EL students to diverse perspectives and thus enrich their learning experiences. Furthermore, by drawing connections to EL students' unique backgrounds, the teacher demonstrates and models a respectful attitude to diversity and encourages students' curiosity about multiple ways of doing things (Yoon & Kim, 2012). Finally, teachers should analyze the social act of language, particularly how language use is a critical component of the story sharing process (Fan, 2013).

In order to effectively bring these stories into the classroom, teachers should see students as knowledge producers and linguistic experts (Bucholtz et al., 2014). All too often in educational contexts, the teacher is the knower and the students are seen as lacking in knowledge (Freire, 2000). Pedagogy viewing students as experts and capitalizing on their funds of knowledge, supports "young people's construction of powerful identities for themselves, identities in which both their academic aspirations and their linguistic and cultural background have an equal place and are mutually reinforcing rather than conflicting" (Bucholtz et al., 2014, p. 149).

Incorporating students' backgrounds and expertise into a classroom effectively is a challenging task. All students have multi-layered identities, and there is no one-size-fits-all method for fostering an inclusive classroom community and learning from students' perspectives (Ciechanowski, 2013). Atkinson and Sohn (2013) explored this notion through their case study of one particular student's cultural identity, uncovering the reality that there was no coherent whole. Instead, within this one student there were multiple, intersecting, often conflicted identities. The concept of intersectionality

is complex, as it underscores the impossibility of exploring one identity alone (Hill Collins & Bilge, 2016); as the social world is explored and analyzed, one must account for the many facets of identity (Crenshaw, 1991). For example, as teachers approach their students, they have to think of the intersecting dynamics of race, language, gender, class, sexual orientation and so on. Teachers have to bring their awareness not only of interlocking identity categories, but also of systems of oppression (racism, heterosexism, etc.) into their teaching practice, combating the impact that powerful groups have over marginalized groups (Adams, Bell, & Griffin, 2007). Furthermore, an essential understanding that teachers have to develop is that such systems of oppression are not static but rather are multidimensional, variable, dynamic and continuously evolving in line with the changing society and developments in the human history (Bell, 2007).

Thus, learning about and discussing linguistic and cultural stories is insufficient; they have to be viewed through power structures, which can help foster a rich and safe learning space among teachers and students. Nieto and Bode (2018) explain that language and cultural issues have to be critically examined through a lens of equity and power in order to bring about change and sustainable learning opportunities. The development of the *Social Justice Standards* as part of the Southern Poverty Law Center' (SPLC) Teaching Tolerance project presents a solid attempt to integrate the critical principles of social justice into K-12 education. The standards used in this article to frame the rubric for the lesson plan analysis were developed based on the four goals of anti-bias education, centering on (1) *identity*, (2) *diversity*, (3) *justice* and (4) *action* (Derman-Sparks & Olsen Edwards, 2010). Students need to learn about who they are and what makes up their social/group identities (*identity*). They also need to learn about one another, creating connections (*diversity*). Then, they develop language to recognize and describe injustice (*justice*) and get empowered to act against prejudice and discrimination (*action*).

Guiding students on a path from self-awareness to collaborative learning and then ultimately action in the face of injustice requires a teacher to be self-aware and self-reflective of his/her identities and how they impact their students. In addition to knowing themselves, teachers have to find opportunities to investigate social dilemmas with their students. Finding these opportunities to connect lesson content to social dilemmas that are relevant to students is a strategic balancing act requiring teachers to "become jugglers to navigate across multiple objectives within a lesson and explorers to embrace opportunities to question and investigate social dilemmas with students" (Ciechanowski, 2013, p. 20). A classroom community needs to have peaceful relationships that are actively nurtured. As students are engaged in conversations about peace building, their critical thinking and conflict resolution skills develop. Jakar and Milofsky (2016) state, "The basic concepts and skills in conflict transformation and peacebuilding deal with reducing prejudice, building relationships, communicating effectively, and using negotiation to manage disagreements" (p. 44). In order to do this, Jakar and Milofsky (2016) recommend integrating multiple perspectives, teaching dialogue, engaging students in interaction, sharing stories and empowering students.

Delpit (1995; 2006) discusses the inherent power imbalances in our classrooms, and how students from marginalized communities often get the message they are failing and struggling in school, when in fact the schools are the ones that need to change to better support these students. Delpit (1995; 2006) explains that teachers need to know and connect with the cultures of their students and also teach them the skills they need in order to navigate powerful discourses in society. In other words, teaching students to write requires both practicing explicit language skills for successful communication in powerful discourses and honoring one's home cultural and linguistic knowledge. This combination will equip students to become savvy navigating different genres and code switching, both critical for their academic and personal success. De Jong and Harper (2005) explain that teachers of all students, and in particular ELs, need to not only have an understanding of students' linguistic

and cultural backgrounds, they also need to have critical linguistic awareness that will allow them to create instructional activities to best meet ELs' needs and include them in the learning process.

Tomita and Spada (2013) explain that form-focused activities in language instruction provide more opportunities for learners to successfully communicate in second language practice. If they know what specific structure they are practicing, they can use that to explain themselves and participate in classroom dialogue. Simply expecting language learners to engage in dialogue in a classroom, without providing them with linguistic structures to join the conversation or add to a point is unfair to ELs and others who do not yet have the language for joining such a conversation. Similarly, intercultural contact has to be actively facilitated and structured to allow intercultural learning (Bennett, 2009). Just putting students in groups, without providing them specific activities and learning tasks, does not mean that they will learn from one another and in some cases, they will actually increase their prejudice across differences in those situations (Spencer-Rodgers & McGovern, 2002).

Thus, teachers must integrate students' backgrounds and story-sharing into their classroom communities, while also tackling power imbalances and guiding students to question the status quo and take action in the face of injustice. In order to do this, they must honor where the students are from, bringing their languages, discourses and funds of knowledge into the classroom dialogue. Furthermore, they should explicitly teach them how to navigate powerful discourses and participate actively in intercultural communication. There is not one method to do this, as the form of education for social justice is not predetermined; instead, it has to emerge from the context and community members' (students') priorities (Bucholtz et al., 2014).

Study Rationale

While social justice principles are addressed in TESOL preparation courses and students are expected to integrate these principles into their lesson plans and instruction, research exploring educational experiences and effective teaching practices for ELs demonstrates that teachers primarily focus on immediate language and content related needs and requirements, such as identifying and meeting language objectives, integrating four language skills, teaching academic vocabulary and others (Bartolomé, 2003; Dutro, 2005; Mize & Dantas-Whitney, 2007; Sensoy & DiAngelo, 2012). It is less common to focus specifically on social justice principles and how effectively they are embedded unless the goal of the lesson is to promote inclusion and social justice. This lack of attention to practical applications of social justice principles may be due to a number of reasons, such as teachers dealing with an overwhelming amount of crucial components in lesson planning and teaching; lacking a clear understanding and repertoires of practical activities grounded in social justice; and failing to recognize the need to focus on social justice principles across various subject areas.

There has been some research into teacher preparation programs and how they integrate culturally responsive and social justice pedagogy into their curricula, impacting teachers' self-awareness and social justice knowledge and classroom practices (Ruffin, 2016; Thieman, 2016), but these studies do not address lesson content. This current study, in contrast, examines the integration of social justice principles in lesson plans developed by teachers who did not have extensive social justice curricula built explicitly into their teacher preparation courses. The focus on lesson plans is a narrow lens that will allow for a deeper discussion of the components of social justice education not only in TESOL preparation, but across a wide variety of teacher preparation programs. This document analysis, explained below, explored the understanding of social justice principles and readiness to enact them in lessons among pre-service and in-service teachers, who are enrolled in the TESOL graduate degree programs. Furthermore, the study findings were used as a starting point to develop specific suggestions about the integration of social justice principles into teaching.

Methodology

Data Sources

The current study reports findings from a document analysis of 50 lesson plans developed by pre-service and in-service English as a second language (ESL) teachers who were pursuing a TESOL graduate degree (Initial License) at a university in the Northeast of the United States at the time of the study. The lesson plans were submitted as part of assignments in two courses: reading and writing for ELs and second language teaching and learning. Students usually take these courses during the first year of their TESOL program. The lesson plan assignment in the reading and writing course asked students to both integrate specific strategies for teaching reading and writing skills acquired in the course and ensure that speaking and listening skills were embedded in the lesson. Students submitted two of such lesson plans over the course of the semester: one lesson plan in the middle of the course semester and the other one towards the completion of the course. Lesson plans from both assignments were analyzed in this study. The lesson plan assignment in the second language teaching and learning course was due at the end of the semester.

It is noteworthy to mention that the lesson plan assignment in either of the courses did not explicitly ask students to integrate social justice principles. Social justice, however, was addressed on different occasions in the courses as part of classroom activities and in course materials. Moreover, students had multiple opportunities to practice analyzing, critiquing and reflecting on lesson plans as part of the instructor's demonstrations, groups work and whole class activities. During such activities, the course instructor drew students' attention to social justice issues in delivering the analyzed lessons, and opportunities for the integrations of social justice principles into the lesson plans were discussed.

In addition to course-specific directions for the lesson plan assignments, such as the inclusion of language objectives, the balance of four language skills and other elements, the expectation for the lesson plans was to ensure that the they were authentic and included principles of good teaching practices acquired in TESOL courses and through experience. Furthermore, students were encouraged to provide thorough descriptions of lesson activities detailing both student and teacher actions in the lesson. As a final component of the lesson plan assignment, students were asked to write a commentary in which they reflected on their lessons self-evaluating their effectiveness and addressing additional considerations that they kept in mind for the lesson plan delivery. This commentary was included into the data analysis together with lesson plans and served as an additional data source that helped the researchers grasp important nuances about students' understanding of social justice and readiness to enact it in lessons.

Data Analysis

The document or documentary analysis method was utilized to examine lesson plans and construct understanding about students' readiness to effectively enact social justice principles in their teaching. This qualitative research technique is used to analyze documents that carry information about the phenomenon or issue under investigation (Bailey, 1994). Payne and Payne (2004) referred to document analysis as a systematic approach to identifying, analyzing and interpreting the content and context of documents.

Both researchers began the analysis with independent close reading and critical examination of the lesson components. This process was accompanied with thematic coding based on the rubric designed by the researchers and described below. The second phase of the analysis included

calibration, which allowed the researchers to compare and validate emerging themes in order to increase trustworthiness of the data. Similar to the independent analysis phase, the rubric was used during the collaborative data analysis phase.

During the independent and collaborative data analysis process, researchers were aware of the existence of possible but unintentional biases and positionality on their end that could impact their own interpretation of the lessons (Bourke, 2014). In order to minimize the potential biases, both researchers participated actively in a variety of social justice workshops among colleagues and engaged in identity activities with their students as part of the social justice curricula. They were both committed to social justice for all and were critical of curricula that marginalizes and excludes. Both researchers had disparate experiences with oppression personally, which contributed to their increased awareness of the ways in which these experiences might inform of critical realities and shield them from others. The rubric described below was developed in line with the idea of reducing the potential impact of biases and positionality of the researchers.

Rubric for Lesson Plan Analysis

The rubric, specifically designed for this study to analyze lesson plans for the inclusion of social justice oriented practices, is aligned with the *Social Justice Standards: The Teaching Tolerance Anti-Biases Framework*, and reflects four key domains: Identity, Diversity, Justice and Action. This framework was developed as part of a project of the Southern Poverty Law Center (1991) known as Teaching Tolerance, and is aimed at promoting K-12 education free of prejudice. The four domains, Identity, Diversity, Justice and Action, comprise twenty anchor standards, which can be used as a guide for curriculum development at every grade level. This comprehensive framework grounded in many years of work promoting anti-bias education and social justice was chosen for the analysis because it encompasses the crucial facets of social justice addressed in this article and specifies behaviors, skills and knowledge that students need to develop over the course of K-12 education.

The rubric comprises nine questions presented below with pertinent domains from the Teaching Tolerance Framework in parentheses:

1. Are multiple perspectives welcomed and respected? (Diversity; Identity)
2. Is curiosity about diversity encouraged? (Diversity)
3. Are self-reflection practices integrated into the lesson? (Identity)
4. Is story sharing included in the lesson? (Identity; Action)
5. Is dialogue part of the instruction? Are students taught explicitly how to engage in an empathetic, open-minded dialogue? (Diversity; Action)
6. Is there any indication of insensitivity to diversity or hidden stereotypes/bias in the lesson? (Justice)
7. Are students' prior learning experiences and background knowledge integrated? (Identity; Diversity). Is there an in-depth exploration of background knowledge/culture (avoiding oversimplification)?
8. Are form-focused instruction and learning strategies included into the lesson in order to provide access to academic learning communities and foster positive identity development? (Identity; Justice)
9. Does the lesson promote action against injustice? (Justice; Action)

Findings and Discussion

The analysis of the data yielded four major themes: (1) Missed Opportunity for Connection; (2) Unstructured Dialogue; (3) Intention to Meet Linguistic Needs; and (4) Lack of Modeling of Intercultural Practices.

(1) Missed Opportunity for Connection

It was found through the coding process that pre-service and in-service ESL teachers were not always successful at connecting the lesson content and language that they planned to teach to students' backgrounds and prior learning. This theme is aligned with the Diversity, Identity and Justice domains of the *Social Justice Standards* (Teaching Tolerance, 2016) and was identified in the analysis based on the following questions in the rubric:

1. Are multiple perspectives welcomed and respected? (Diversity; Identity)
2. Is curiosity about diversity encouraged? (Diversity)
6. Is there any indication of insensitivity to diversity or hidden stereotypes/bias in the lesson? (Justice)
7. Are students' prior learning experiences and background knowledge integrated? (Identity; Diversity). Is there an in-depth exploration of background knowledge/culture (avoiding oversimplification)?

Most of the lesson plans included "missed" opportunities to introduce multiple perspectives and provide connections to ELs' backgrounds in the activator section, main body and/or homework assignments. The activator was a required component in the lesson plan assignment and was intended to encourage students to think about ways to connect new concept to students' lives and experiences. However, the analysis revealed that such connections, when provided, were usually shallow and one-sided reflecting only the dominant cultural perspective. For example, in one lesson plan designed for a group of 20 students including four students from Cape Verde, the teacher introduced weather patterns for different seasons as an example of "fact" versus "opinion" sentences. One example of a fact sentence was "The weather is 25F in the winter". The lesson activator and the main body did not include any indication that the temperature is different in the winter depending on the region and/or country. There was a missed opportunity to ask Cape Verdean students to share what the weather is like in Cape Verde in the winter, which contrasts with the temperature in the Northeast of the U.S. As is evident, this example did not only fail to provide connections to EL students' experiences and background knowledge but also could create confusion among the Cape Verdean students who might have a different concept of the winter season, especially if they were newcomers. This could impact students' understanding of the "fact" concept, which was the topic of the lesson.

Another example that supports this theme was found in a lesson on school rules for 16 sixth-grade students among which there were six ELs. The lesson included a handout listing common school rules in the U.S.; however, it did not include any discussion of how school rules vary in different countries and contexts. Provided there are six ELs, such a discussion would create an opportunity for meaningful and engaging learning for all students in the classroom and would foster a safe environment welcoming multiple perspectives and ways of doing things. In order to engage students further, the teacher could create an activity asking students to identify similar and different rules among cultures and discuss them as a whole class. Finally, the teacher could allow students to choose rules from the ELs' cultures that they could add to the list and thus further reinforce the inclusion.

Such practices could also encourage curiosity among non-ELs and introduce them to a multiplicity of perspectives.

In each of the examples above, it is important to engage students in activities that extend beyond just an introduction and superficial discussion of diverse perspectives. Instead, the teacher should encourage students' in-depth, critical reflection leading them to develop a non-judgmental, respectful attitude to diversity, ensuring that all voices are heard.

(2)   Unstructured Dialogue

The theme of Unstructured Dialogue is closely aligned with the finding discussed above and presents an important aspect of social justice oriented teaching. Interaction plays a crucial role in a classroom of ELs for a number of reasons. First, it provides meaningful opportunities for students to practice oral language skills and jointly develop an understanding about the lesson topic in the Zone of Proximal Development, which is found to have a significant impact on language attainment including all four language skills (Lantolf & Thorne, 2007). Second, structured classroom communication motivates cooperative learning and negotiation of meaning while providing all students with an opportunity to explore and practice respectful dialogue that welcomes various perspectives.

It is imperative to explicitly teach norms and rules of effective and productive classroom communication to all students who may have various degrees of familiarity with expectations for engaging and sustaining a respectful dialogue in a U.S. school setting. These differences may result in an inadequate participation in classroom oral exchange activities. For many students, this different cultural understanding is accompanied by a lack of English language skills necessary to engage in dynamic communicative activities with native speaking peers, creating an additional obstacle for dialogue opportunities.

This theme falls into with the Diversity and Action domains of *Social Justice Standards* (Teaching Tolerance, 2016). The question in the Rubric that helped identify this theme was:

> 5. Is dialogue part of the instruction? Are students taught explicitly how to engage in an empathetic, open-minded dialogue?

All analyzed lessons included some indication of interaction opportunities. Nevertheless, oral communication activities were not always sufficiently structured, and essential rules and guidelines for engaging in a respectful dialogue were not explicitly addressed, modeled and practiced. In one lesson plan, a group discussion was included to serve as a prerequisite for the next step – writing an opinion essay. Each student was expected to express an opinion about his/her favorite season and explain why he/she thought that way. This phase was critical for developing students' understanding of the concept of opinion and how opinions, unlike facts, can vary among individuals. However, the lesson failed to include explicit explanation of how students should operate within a group to ensure that everyone felt comfortable expressing opinions. Such an explanation was particularly necessary in the context of the classroom where the lesson was taught due to a large proportion of ELs - 12 students out of 16 students were ELs.

In another lesson, 20 fourth graders including five ELs were expected to work in groups and discuss food color, texture, smell and taste. As part of this group activity, students needed to share which food they liked or disliked and why. Similar to the previously discussed lesson, there was a lack of teachers' modeling of ground rules of participating in a dialogue of this sort. Such modeling was essential in order to introduce variations among individual preferences and perceptions of food flavors and smells, which may drastically differ across cultures. What is considered spicy in one culture may be viewed as neutral in another culture. Such variations can lead to misunderstandings and reduce

productivity if the above rules are not in place. Without teaching conventions of respectful dialogue, EL students sharing knowledge based on their experiences may put themselves at risk and feel inadequate.

Yet another lesson was designed for a classroom of 18 seventh grade students among which four students were ELs. The lesson included an activity asking students to read each other's paragraphs in which they shared their interpretations of the moral conveyed in a story read earlier. As part of this activity, students needed to express agreement or disagreement with their peers' views and defend their positions. The teacher planned to scaffold reading of the story by introducing and explaining key vocabulary and engaging students in a discussion of various parts of the story. She also provided necessary linguistic supports with writing a paragraph to ELs who were at lower levels of English development. However, the lesson did not address expectations for the oral exchange activity encouraging students to encounter diverse opinions and respond to them appropriately. The insufficient structure of this communicative activity could put EL students at risk as they might lack conversational tools and linguistic skills necessary to sustain the discussion.

One important aspect of group interaction activities that is worth specific attention in a classroom of culturally and linguistically diverse students is grouping structures. The data analysis revealed that only some (roughly 1/3) lesson plans specified criteria for grouping configurations. Placing ELs in appropriate groups informed by students' levels of English proficiency, cultural backgrounds, personality characteristics and other criteria, is important as it can significantly affect the outcomes of a discussion activity in terms of student learning and the level of comfort working in a group.

(3) Intention to Meet Linguistic Needs

The third major finding is Intention to Meet Linguistic Needs. In order for ELs to be able to engage in higher-order thinking and fully participate in all class activities, it is critical that their language development is supported (Echevarría & Graves, 2010).

This theme relates to the Identity and Justice domains of the *Social Justice Standards* (Teaching Tolerance, 2016). Students whose access to academic learning and the classroom community is restricted as a result of inadequate language instruction will be unable to develop positive social identities, and their membership in the learning community will be restricted. If learners do not have necessary linguistic skills to participate in learning tasks, their academic progress will be at risk. The question in the rubric that addressed this area is:

8. Are form-focused instruction and learning strategies included into the lesson in order to provide access to academic learning communities and foster positive identity development? (Identity; Justice)

Many of the lessons explored in this study included intent to differentiate based on English proficiency levels but did not have a clear structure for enacting the language instruction and leveled activities. The differentiation plan and supports for language learning were usually explained in the lesson context section and then referenced throughout the lesson. Some specific strategies and supports to appropriately teach English language were integrated into all of the lessons. For example, differentiated sentence frames, graphic organizers, group work activities, vocabulary instruction, differentiated writing activities and language assessments were included at various points during the lesson activities.

While there were many strategies and supports identified, the effectiveness of language instruction and activities was often unclear. For example, there was an intent to teach vocabulary, but instead of describing the specific method of vocabulary instruction, there was often just a general statement about the need to teach the words. For example, in one of the lessons, the teacher explained that her "lesson starts out in the domain of oral language, with students reviewing vocabulary and working with partners to verbalize what they've already learned about animal homes" without clearly

explaining how this vocabulary instruction would be structured. In order to meet the linguistic needs of ELs and thus ensure that they are able to participate fully in lessons, carefully structured vocabulary instruction that is differentiated according to proficiency levels is essential.

Furthermore, while all of the lessons included language objectives, many of them were not effectively structured and it was unclear whether there were activities that would meet and assess those language objectives as the lesson unfolded. For example in one lesson, the language objective was, "I will be able to list reasons why I like my favorite season", but it was not clear whether the list would be shared orally or in writing. Then as the lesson went on, the process of creating a list of reasons for a favorite season was not clearly defined in the lesson activities and it did not seem to be assessed at any point. The assessment was instead focused on opinion writing and whether there were complete sentences and adequate reasons. In order to meet ELs' linguistic needs, lessons need to have effective language objectives that are met through lesson activities; if this component is missing ELs' ability to participate fully in the lesson will limit their access to curriculum.

Lastly, there were a variety of instances where learning strategies were mentioned, but it was not apparent whether the strategies were taught explicitly during the lesson. For example, as previously mentioned, while pair and group work activities were integrated into most lessons, the interaction rules and norms as well as the linguistic structures necessary for participation were often inadequately explained and structured. For example, in one lesson the teacher candidate explained that in the lesson activator, "The other students will be asked to discuss with a partner what they could do if they had those materials shown in the picture." Not only was there no mention of how respectful communication would be fostered, but also there was inadequate support for lower levels of ELs to engage in the interaction.

Another example of inadequate explanation of a learning strategy refers to the integration of graphic organizers in the lessons. Graphic organizers were often mentioned as a tool for students to organize their information, but were not fully explicated to students. If students do not have experience with a particular graphic organizer (eg. a Venn diagram as was the case in a few lessons), then they have to first learn about how to use the graphic organizer in order for it to be helpful in their language and content learning. If graphic organizers have already been explained to students, then that should be mentioned in the lesson.

(4)   Intercultural Practices Not Modelled

The last major theme aligns with the Identity and Action domains of the *Social Justice Standards* (Teaching Tolerance, 2016) and was generated based on the following questions from the rubric:

3.   Are self-reflection practices integrated into the lesson? (Identity)
4.   Is story sharing included in the lesson? (Identity; Action)
5.   /…/ Are students taught explicitly how to engage in a empathetic, open-minded dialogue? (Diversity; Action)
10.   Does the lesson promote action against injustice? (Justice; Action)

The lesson plan document analysis demonstrated that the teachers did not model intercultural practices across all the lesson plans. In addition to missing opportunities to connect with students' backgrounds, as was described above, the teachers did not exhibit self-reflection, curiosity, openness and knowledge of multiple perspectives themselves in order to show students examples of what these practices looked like. It is possible that the teachers creating the lessons had yet to acquire intercultural experiences themselves and thus were not adept at bringing in multiple perspectives or modeling intercultural skills, underscoring the need to foster the development of critical intercultural

competence in teacher education courses. Another possibility for the inadequate modelling of intercultural practices or lack thereof in lessons is teachers' tendency to assume that intercultural competence develops naturally and is not worth lesson time and attention, or it is simply not part of the curriculum. In fact, it is known that intercultural competence development is a lifelong process that requires explicit teaching and scaffolding stimulating ongoing critical self-reflection and analysis (Bennett, 2009).

Thus, the teacher needs to enact intercultural attitudes, knowledge, skills and behaviors himself/herself in order to be an exemplar of how an intercultural communicator should act like (Deardorff, 2006). For example, when the teacher introduces topics, there is almost always an opportunity to model and show awareness of multiple perspectives. At the beginning of one lesson, the teacher asked the students to discuss collectable items. Prior to having the students interact, the teacher could model self-reflection, explaining his/her thought process at this point, and then review examples that show awareness of multiple perspectives. The teacher could bring a cultural artifact to share with the class (e.g. international coins) that he/she collected and then explain how a friend or a classmate from another cultural background collected something completely different (e.g. tea cups), demonstrating the values, beliefs and traditions unique to that culture. Ideally, the teacher could share the actual objects or at least pictures of the objects. Then, the discussion among the students could follow from the teachers' examples. Rarely did the teacher share his/her own examples that exhibited an understanding and curiosity of multiple perspectives.

In addition to not sharing his/her own examples, the teacher expected the students to be the ones to bring in their diverse ideas and engage with one another without adequate scaffolding, which could inhibit students' exploration of multiple perspectives. It was the students' responsibility to structure this process, rather than having the teacher facilitate the learning experience for students. In some lessons, the teacher asked students to share their own stories, but did not explicitly structure the conversation to elicit a respectful, in-depth exploration of background knowledge. In one lesson, a teacher described how the group sharing would occur by stating, "They will one by one share one of their brainstormed words by standing up and acting it out, a fun way to get moving and share their ideas with the class, as well as providing the possibility for culturally diverse students to show a quick glimpse into something children do or experience during a certain season where they are from." This fun "quick glimpse" was not a sufficient explanation of a lesson activity that could foster an inclusive environment and a productive conversation across cultures.

Furthermore, another way to model intercultural competence is to ask follow up questions, respond to the nuances of students' ideas and design various opportunities for them to share their work. Throughout the lessons, there were many points when the teacher would ask the students to do an activity and then share out to the whole group, without modelling respectful communication prior to or during the activity. The teacher could first ask follow-up questions to the students, demonstrating what intercultural curiosity looks like in practice. Then, the teacher could encourage the students to ask one another similar questions before sharing with the whole group.

Finally, none of the lesson plans included examples or practices of how students could plan and carry out individual and collective action against prejudice and injustice. It was not surprising that the action domain was missing as the other components of social justice, such as self-reflection, awareness about diversity and respect, which serve as prerequisite steps for action, were not emphasized. These elements must be properly scaffolded in order to guide students in developing an in-depth and genuine understanding about critical nuances of diversity, which can enable students to recognize their responsibility to stand up to injustice and develop an active stance with regard to exclusion, bias and oppression in the local and global community. Such complex competencies are not a matter of one lesson or unit. Instead, it may take months or rather years to ingrain these principles in students' repertoires. Therefore, it is essential to introduce students to issues of diversity

early in their academic experiences and reinforce crucial social justice aspects throughout their school years.

Limitations and Future Directions

There were a number of limitations in the study methodology that need to be acknowledged. Lesson plans were the only source of data that was used in the analysis of teachers' understanding of social justice principles and readiness to enact them in the classroom. It is worth mentioning, however, that the lesson plan assignment specifically asked to provide thorough descriptions of lesson activities including student and teacher roles. In addition to the lesson plan, teachers wrote a commentary reflecting on the effectiveness of their lesson plans, potential challenges and other considerations.

The depth of the lesson description and the commentary were valuable for the analysis in the study of this scope. Nevertheless, researchers acknowledge the need to conduct a more comprehensive investigation of this issue in order to understand potential weaknesses in teacher education programs with regard to preparing effective teachers who are ready to enact pedagogy for social justice. Thus, in addition to lesson plans, actual classroom observations, in which these lesson plans are used, need to be included. Such observations could yield interesting findings about critical intercultural issues in the classroom.

Furthermore, teachers' reflection collected through interviews or focus groups could contribute to understanding their thinking and decision making process when they plan a lesson and enact social justice principles in their teaching. Researchers did not have data about the participants' identity groups. Knowing about the groups students identify with and their experiences of oppression might have contributed to the ways in which researchers were able to analyze their lesson plans. For a future research study, exploring how teachers' identities could inform their integration of social justice standards would be a worthwhile investigation.

Finally, it was beyond the scope of this study to look at similar teacher preparation programs at other universities in the U.S. However, in order to construct a nuanced understanding about teachers' readiness to serve as social justice advocates in American schools as well as gauge the quality of teacher education programs with this regard across the U.S., it is necessary to examine similar programs at colleges and universities in different parts of the country.

Conclusion and Recommendations

The document analysis suggests that a variety of teacher education programs, such as early literacy, elementary and secondary education, TESOL, special education, educational leadership and other teacher preparation programs, should include social justice orientations as required components in the curriculum. The inclusion of this sort must go beyond mere discussion of issues of injustice and the importance of welcoming diversity in the classroom. Rather it must be structured in a way that teachers develop a thorough understanding of social justice principles and acquire tools necessary for enacting them in their practices. Such competencies can evolve if teachers are engaged and guided through activities stimulating critical thinking, reflection, analysis of issues of diversity and implementation of social justice principles across teaching contexts.

Consistent with prominent models of intercultural competence (Deadorff, 2006) and critical intercultural communication (Halualani & Nakayama, 2010), self-reflection practices leading teachers to the awareness of their own identities and readiness to act as social justice agents must be included as they become educators committed to social justice. Such reflection activities must be systematic, in-depth, deliberately planned and relevant to teacher' experiences. When teachers are aware of their own identities and orientations, and their role in structural and systemic inequity, they can relate to

others (their students), comparing and contrasting their characteristics, and embracing the multiplicity of perspectives. With this foundation in place, it is essential to present teacher candidates with case studies, real-life examples of individuals or contexts and experiential activities in which identities are put at risk due to blatant or hidden injustice.

Another crucial recommendation for programs grounded in social justice, and supported by the findings of this study, is introducing practical strategies that instructors can use across pedagogical contexts to create socially just education climates. One fundamental practical strategy that has been found to be crucial for teaching culturally and linguistically diverse students is including form and meaning focused instruction across all content areas, ensuring equitable access to the academic curricula (Short & Echevarría, 2016; Tomita & Spada, 2013). Such an explicit focus on linguistic demands is essential as it lays the foundation for a socially just learning environment.

Among other practical strategies is modeling a structured and respectful classroom dialogue that welcomes diverse opinions and promotes collaborative learning. Instructors in teacher preparation programs need to structure discussions in a way that all students feel heard and know how to listen to others as well as possess necessary group discussion skills. Some of such skills are: do not interrupt, avoid judgment, use appropriate academic language when disagreeing, do not monopolize the discussion, etc. In K-12 teacher preparation, it is essential that teachers learn and practice how to enact such strategies through lesson plan assignments, practicum activities, and community service across the program curriculum. A special emphasis in teacher preparation needs to be placed on promoting individual and collective action against injustice, the component that was not identified in the findings of this study.

The research implications suggested above extend beyond teacher education programs to a variety of academic disciplines. All teaching and learning situations regardless of the academic field include opportunities for faculty and students to interact across differences and enact social justice principles. In addition to linguistic and cultural diversity discussed in this article, the same social justice principles should be applied with regard to other diversity characteristics that comprise an individual identity, such as race, ethnicity, religion, disability, gender, among others, which are present on every university and college campus. In such contexts, it is critical that a diversity approach be replaced by a social justice approach which underscores the role that inequities play in all interactions, particularly in teaching and learning contexts where there is a heightened amount of communication on critical topics (Adams & Zuniga, 2016).

Based on our findings, all faculty across academic discipline should engage in self-reflection practices and explore their own identities as a fundamental initial step in understanding their own roles and responsibilities in the multicultural context. Such self-reflection will also enhance faculty's understanding of power imbalances in the classroom, helping them facilitate equitable classroom dynamics (Bell, Goodman & Varghese, 2016). Furthermore, they should strive to acquire practical strategies that they can use to enact social justice principles in the classroom context. Finally, faculty across various disciplines should model respectful dialogue, freedom of opinions and safe collaborative environment on a daily basis in their classes. In order to make teaching and learning more equitable, providing opportunities for all students to be included and to be able to express themselves, it is essential that instructors connect with their students, provide structure to intercultural dialogue and model intercultural practices. For faculty in higher education, there should be enhanced professional development opportunities that allow for sharing of ideas among faculty and an opportunity to explore and practice these inclusive strategies. Moreover, the university climate in general must be conducive to promoting respect to cultural diversity and implementing such social justice driven practices.

References

Adams, M., & Bell, L. (Eds.). (2016). *Teaching for diversity and social justice* (3rd ed.). New York, NY: Routledge.

Adams, M., & Zuniga, X. (2016). Getting started: Core concepts for social justice education. In M. Adams & L. Bell (Eds.), *Teaching for diversity and social justice* (pp. 95–130). New York, NY: Routledge.

Atkinson, D., & Sohn, J. (2013). Culture from the bottom up. *TESOL Quarterly, 47*(4), 669-693. https://doi:org/10.1002/tesq.104

Ayers, W., Quinn, T., & Stovall, D. (2009). *Handbook of social justice in education*. New York: Routledge.

Bartolomé, L. (2003). Beyond the methods fetish: Toward a humanizing pedagogy. In Darder, A., Baltodano, M., & Torres, R. (Eds.). *The critical pedagogy reader* (pp. 173-195). NY: Routledge. https://doi.org/10.17763/haer.64.2.58q5m5744t325730

Bailey, K.D. (1994). *Methods of social research.* New York: The Free Press.

Bell, L. A. (2007). Theoretical foundations for social justice education. In M. Adams, L. A. Bell, & P. Griffin (Eds.), *Teaching for diversity and social justice* (pp. 1-14). New York, NY, US: Routledge/Taylor & Francis Group.

Bell, L., Goodman, D., & Varghese, R. (2016). Critical self-knowledge for social justice educators. In M. Adams & L. Bell (Eds.), *Teaching for diversity and social justice* (pp. 397-418). New York, NY: Routledge.

Bennett, J. M. (2009). Cultivating intercultural competence: A process perspective. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 121-140). Thousand Oaks, CA: SAGE Publications, Inc.

Bourke, B. (2014). Positionality: Reflecting on the research process. *The Qualitative Report, 19*(33).

Bucholtz, M., Lopez, A., Mojarro, A., Skapoulli, E., VanderStouwe, C., & Warner-Garcia, S. (2014). Sociolinguistic justice in the schools: Student researchers as linguistic experts. *Language and Linguistics Compass, 8*, 144-157. https://doi.org/10.1111/lnc3

Canagarajah, A. S. (2014). Theorizing a competence for translingual practice at the contact zone. In S. May (Ed.), *The multilingual turn: Implications for SLA, TESOL and bilingual education* (pp. 78-102). New York: Routledge.

Ciechanowski, K. (2013). Beyond one-size-fits-all ELD frameworks: Bringing English learner's lives and social justice to the center of K-12 instruction. *ORTESOL Journal, 30*, 15-20.

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review, 43*(6), 1241-1299. doi:10.2307/1229039

Darling-Hammond, L. (2010). *The flat world: How America's commitment to education will determine our future.* New York: Teachers College Press.

de Jong, E.J., & C.A. Harper (2005). Preparing mainstream teachers for English language learners: Is being a good teacher good enough? *Teacher Education Quarterly, 32*(2), 101–24.

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States. *Journal of Studies in International Education, 10*, 241-266. doi:10.1177/1028315306287002

Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom.* New York: New Press.

Delpit, L. (2006). Lessons from teachers. *Journal of Teacher Education, 57*(3), 220-231.

doi:10.1177/0022487105285966

Derman-Sparks, L., & Edwards, O. J. (2010). *Anti-bias education for young children and ourselves.* Washington, DC: NAEYC

Dutro, S. (2005). *A focused approach to systematic ELD: A handbook for K-6 teachers.* San Marcos, CA: E.L.Achieve.

Echevarría, J., & Graves, A. (2010). *Sheltered content instruction: Teaching English learners with diverse abilities* (4th ed.). Boston, MA: Allyn & Bacon.

Fan, Y. (2013). Every teacher is a language teacher. *International Journal of Community Research and Engagement, 6,* 77-92.

Freire, P. (2000). *Pedagogy of the oppressed* (30th anniversary ed.). New York: Continuum.

Good, & S. M. Putman (Eds.), *Handbook of research on professional development for quality teaching and learning* (pp. 409-426). Hershey, PA: IGI Global.

Hill Collins, P., & Bilge, S. (2016). *Intersectionality (Key concepts).* Cambridge, UK: Polity Press.

Halualani, R. T. & Nakayama, T. K. (2010). Critical intercultural communication studies: At a crossroads. In R.T. Halualani & T.K. Nakayama (Eds.), *The handbook of critical intercultural communication* (pp. 1-16). Malden, MA: Blackwell Publishing Ltd.

Jakar, V. S. & Milofsky, A. (2016). Bringing peacebuilding into the English language classroom. In C. Hastings & L. Jacob, L. (Eds.), *Social justice in English language teaching* (pp. 39-46). Alexandria, VA: TESOL Press.

Lantolf, J. & Thorne, S. L. (2007). Sociocultural theory and second language learning. In. B. van Patten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum.

May, S. (Ed.). (2014). *The multilingual turn: Implications for SLA, TESOL and bilingual Education.* New York: Routledge, Taylor & Francis Group.

Mize, K. & Dantas-Whitney, M. (2007). English language development in K-12 settings: Principles, cautions, and effective models. *ORTESOL Journal, 25*, 17-24.

Nieto, S., & Bode, P. (2018). *Affirming diversity: The sociopolitical context of multicultural Education* (7th edition. ed.). New York: Pearson.

Norton, B. (2014). Identity, literacy and the multilingual classroom. In S. May (Ed.), *The multilingual turn: Implications for SLA, TESOL and bilingual education* (pp. 103-121). New York: Routledge.

Payne, G. & Payne, J. (2004). *Key concepts in social research.* London: Sage Publications.

Peirce, B. N. (1995). Social Identity, investment, and language learning. *TESOL Quarterly*, *29*(1), 9–31. https://doi.org/10.2307/358780

Piller, I. (2016). *Linguistic diversity and social justice: An introduction to applied sociolinguistics.* Oxford University Press.

Ruffin, T. (2016). Equity and inclusion in today's diverse and inclusive 21st century classroom: Fostering culturally responsive pre-service teachers with the tools to provide culturally responsive instruction. In K. Gonzalez & R. Frumkin (Eds.), *Handbook of research on effective communication in culturally diverse classrooms* (pp. 269–283). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-4666-9953-3.ch014

Sensoy, O. & DiAngelo, R. (2012). *Is everyone really equal? An introduction to key concepts in social justice education.* New York, NY: Teachers College.

Short, D. & Echevarría, J. (2016). *Developing academic language using the SIOP model.* New York: Pearson Allyn & Bacon.

Spencer-Rodgers, J., & McGovern, T. (2002). Attitudes toward the culturally different: The role

of intercultural communication barriers, affective responses, consensual stereotypes, and perceived threat. *International Journal of Intercultural Relations*, *26*, 609-631. http://dx.doi.org/10.1016/S0147-1767(02)00038-X

Teaching Tolerance. (2016). *Social justice standards: The teaching tolerance anti-bias framework.* Retrieved from https://www.tolerance.org/sites/default/files/2017-06/TT_Social_Justice_Standards_0.pdf

Thieman, T. (2016). Revising a Teacher education program for diversity and social justice through culturally responsive coursework and professional collaboration. In T. M. Petty,

Tomita, Y., & Spada, N. (2013). Form-focused instruction and learner investment in L2 communication. *The Modern Language Journal, 97*(3), 591-610. https://doi.org/10.1111/j.1540-4781.2013.12031.x

Yoon, B. & Kim, H. K. (2012) (Eds.). *Teachers' roles in second language learning: Classroom applications of sociocultural theory.* Charlotte, NC: Information Age.

Making Connections:
Student-Teacher Rapport in Higher Education Classrooms

Roehl Sybing
University of Massachusetts Amherst
rsybing@umass.edu

Abstract: Educational research asserts the importance of establishing rapport between teachers and their students for the sake of fostering a classroom environment that is conducive to learning. Especially given the disparities in outcomes of university students, it is imperative for educators and policy makers to look at teaching practices in the college classroom as well as policies relevant to teaching and learning in university contexts. This paper reports on an ethnographic study of a college-level academic writing class, centering on how its writing teacher seeks to establish rapport and facilitate understanding with first- and second-year undergraduate students. The findings presented in this paper highlight examples practitioners can examine to validate student knowledge and participation as well as mitigate the effects of differences in identity between teacher and student. This paper closes by inviting discussion and reflection of college-level teachers' practices in the classroom and whether they elicit engagement from students.

Keywords: higher education, classroom interaction, student-teacher rapport

Introduction

Contemporary research has begun to explore the importance of establishing rapport between teachers and students in higher education contexts (Arghode et al., 2017; Estepp & Roberts, 2015; Lee, 2015). Despite this, popular perceptions of higher education do not always center around images of classrooms engaged in discussion and engagement by students. Using the website Google Images, a cursory search of pictures of college classrooms yields images of lecture rooms with large numbers of students passively listening to a teacher who is presumably a master of the relevant course knowledge. The college teacher may not always be expected to connect with their students on an emotional, casual basis, or engage and validate the knowledge that they bring to the classroom. Theory on teaching and learning emphasizing the importance of rapport between teacher and student clashes with the historical traditions of the academy, as well as the feasibility of pedagogies in contexts that serve massive numbers of students.

The traditions of the academy, however, have been shown not to produce outcomes that are equitable to everyone they serve. As sociocultural research emphasizes the differences in culture and identity across the range of diversity encountered in universities in the United States, it is important to recognize that the instructional methods of old may not be suitable for all. With students of color and low-income students graduating at lower rates than white students and high-income students respectively (Musto, 2017; Tate, 2017), it is worth exploring, in addition to reforms of policies, how practitioners in higher education can better accommodate all students.

To that effect, this paper presents an ethnography of a writing teacher promoting a pedagogy that runs counter to that of the perceived typical college lecturer. The narratives presented in this discussion are intended to depict a teacher who values the knowledge and participation of students and mitigates the typical asymmetric power dynamic within the classroom. In so doing, this paper will present a framework that will allow for future research investigating the effects of teacher pedagogy on student-teacher rapport and learner outcomes in higher education contexts.

Historical overview

The need for more dynamic forms of teaching in higher education is apparent in the diversity of students entering college and the outcomes that colleges produce. Verner and Dickinson (1967) were among the first scholars to assert that differences in students in terms of prior education received limited the effectiveness of traditional lecture methods in university education. Moreover, differences in education level are informed both socioeconomic status and ethnicity, particularly in the United States context (Roithmayr, 2014). Recent statistics show that college classrooms in the United States have become more diverse across all cleavages since at least the 1960s in terms of ethnicity, gender, and socioeconomic status, challenging what has traditionally been a white, male, and upper-class higher education context (*National Center for Education Statistics*, 2010). As the college classroom becomes more diverse, it is important to accommodate the growing body of knowledge that diversity brings to higher education.

Diversity in student enrollment, however, has not been matched by diversity in graduation. Graduation rates over six-year periods have highlighted disparities across racial and ethnic lines, with African-American and Hispanic students completing college at a much lower rate than their white and Asian-American counterparts (Tate, 2017). Only 38% of African-American students are shown to complete an undergraduate degree or certificate program within six years, far below the aggregate average of 54.8% of all college students. The income divide yields a starker picture in which only 16% of low-income college students in the United States eventually earn an undergraduate degree compared to 60% of their wealthier peers (Musto, 2017). Absent intervention in the status quo prevailing over higher education, the disparities across race and income (and other cleavages less examined on this issue) are certain to persist and contribute to the social inequities that perpetuate the hegemony of privileged classes at the expense of those at the margins (Roithmayr, 2014).

In examining the recent data connecting student success in higher education to differences in identity and resources, it is unproductive, and perhaps detrimental, to treat all learners in a given classroom in the same manner. For educators at the practitioner level, the implication here is that students at the margins are not served by the status quo in higher education. A different approach, given the disparity in outcomes produced in higher education, is thus required.

Student-teacher rapport

Chickering and Gamson (1987) provided some early definition to pedagogies that transcended assumptions of simple knowledge transfer. In their article about undergraduate teaching, they emphasized that a successful educator:

1. Encourages contacts between students and faculty
2. Develops reciprocity and cooperation among students
3. Uses active learning techniques
4. Gives prompt feedback
5. Emphasizes time on task
6. Communicates high expectations, and
7. Respects diverse talents and ways of learning (p. 3).

This set of principles was intended to address perceived shortcomings that were associated with "[a]pathetic students, illiterate graduates, incompetent teaching, impersonal campuses" (p. 3). Among their most pointed assertions was that learning was not meant to be a passive endeavor

primarily conducted in large lecture rooms with students quietly listening to their professor. Instead, they emphasized that learning was a collaborative process between student and teacher, and even between students. Complementing this belief was the acknowledgment that learners were not uniform, bringing with them diverse and disparate skill sets, ideologies, and dispositions that affected how they expanded their understanding of unfamiliar knowledge. In turn, some of the dispositions that Chickering and Gamson advocate in higher education practitioners, particularly relating to contact between student and teacher and respect for variety of learning styles, require teachers to establish the kind of rapport that is less likely to be established through didactic, almost mechanical pedagogies in classroom settings.

This discussion continues to generate some degree of contention within research on teaching and learning to this day, as evidenced by elements of contemporary literature that still advocate for the effectiveness of lecture-based teaching, one of the mainstays of university education (e.g., Grauer et al., 2008; Taglieri et al., 2017; Thrall et al., 2016). However, such research either has been critiqued for limitations relating to methodology (Thistlethwaite et al., 2012) or, more importantly, highlights the importance of further inquiry into non-lecture methods of teaching (Reimschisel et al., 2017). In either case, even if there is a concession that lecture-based teaching has a place in higher education as an effective pathway to student success, there is an argument to be made that innovations in pedagogy along the lines of teacher rapport with students have yet to be fully identified.

Exploring this aspect of teaching and learning has been made important by the contemporary empirical research on rapport in higher education, which has drawn connections between strategies for building rapport and facilitation of the learning process. Estepp and Roberts (2015), for example, identified discrete strategies for university-level practitioners such as encouraging more interaction in the class, relating personal experiences to the concepts being taught, and proactively expressing respect for their students as pedagogical implications for rapport when indexed to greater learning outcomes. Further research on engagement with student (e.g., Arghode et al., 2017) has also explored how rapport with students can be built through understanding students' perspectives and preferences, which is seen as a means to connect with students in a positive way. Outside of the classroom, Lee's (2015) research with university writing tutors has pointed out the importance of rapport through engaging and conscientious feedback in the writing process in building confidence in students.

Extending this research can potentially yield implications for addressing, at least in part, the wide disparities in outcomes witnessed in higher education when teachers reach out to their students across differences to co-construct knowledge in a manner that students can understand. Scholars (e.g., Colbert, 2010) have drawn connections between the role of pedagogies of active learning in accommodating various learning styles to the development of culturally responsive classrooms, which, in Wlodkowski and Ginsberg's (1995) view, "help students relate lesson content to their own backgrounds" (p. 17), with the assumption that "ignor[ing] student norms of behavior and communication provokes student resistance" (p. 17).

As a result, the study presented in this paper attempts to explore the following question:

RQ1: What pedagogical strategies does a teacher in an academic writing course in a higher university context employ to establish rapport?

Exploration of this research question assumes that rapport between students and their teacher is not a given, particularly because of differences in identity and differences in perceived power. Where students in any formal context are typically considered novices unfamiliar with the content knowledge presented to them, teachers are typically thought of as experts whose job is to transfer necessary knowledge. In addition, institutions of education lend power to their practitioners in allowing them

the ability to assign grades, essentially bestowing the role of gatekeeper onto teachers. This asymmetric power relationship arguably contributes to distance between student and teacher, providing a significant obstacle to establishing rapport built on empathy and trust difficult. The speech and pragmatic acts that practitioners perform under such conditions can provide useful guidance to researchers, educators, and policy makers for examining interaction among participants in higher education.

Theoretical perspective

The framework intended to examine teaching methods in the classroom is founded on the principles for understanding language socialization and the knowledge that language socialization helps to facilitate. Ochs and Schieffelin (2011) explore language socialization as a process in which novices are "socialized" by experts in knowledge, rituals, and discourse. Within this construct, it is overly simplistic to assume that novices lack knowledge of their own; instead, experts are considered experts because the knowledge they have is ratified by the culture in which they exist and interact. This concept can be typically visualized by adults teaching children, or older educators who are masters of content knowledge teaching learners who lack such knowledge. Such teachers are ratified by the culture whether through advanced degrees or teaching certifications, and such ratification perpetuates the need for students to comply in the learning process. In such approaches to teaching and learning, students may not necessarily enter the classroom with zero knowledge, but with a lack of knowledge that is relevant to what educators believe is necessary for post-education life. It is, in such deficit models, necessary for the classroom teacher to transfer the essential knowledge that education assumes students do not have.

       The research presented here looks at a teacher's speech acts and actions that are intended to facilitate understanding of knowledge while negotiating the contributions that learners bring to the classroom. Ochs (2004) provides a useful definition for understanding knowledge and expertise that is the subject for socialization within a context. In her discourse analysis of socialization through rituals taking place at a softball game for children, Ochs makes distinctions among the following elements:

- *categories* (e.g., first base, second base, pitcher, catcher),
- *rules* (e.g., hit the ball and run to first base, avoid being tagged out),
- *expectations* (e.g., being able to swing a bat in a certain manner), and
- *strategies* (e.g., knowing when to run or to stay on base).

       Through this framework, socialization within a community builds and reinforces not only a textbook definition understanding of any particular activity, but an understanding of how one's actions with that activity affects their perceived performance and, thus, their standing within the given community. In this framework, a novice within an activity may have a surface understanding of how an activity is done (via rules), but not a deeper insight in what to do in a dynamic situation (via strategies). More generally, novices may have gaps in insight in any of the four categories, and thus build their insight through experience and effort within their community.

Research site

The research centered on a public university course that focuses on expository writing for academic purposes. The university's writing program organizes the course and requires it for all undergraduate students who do not fulfill the program's writing requirement in advance via tests such as Advanced

Placement exams. Writing teachers in the program are typically hired on an annual, contractual basis, and have various academic backgrounds from education to fine arts. The course is held three times a week, with this section being conducted in the late morning.

The university itself is a major public university with a strong emphasis in research in science, technology, engineering, and math (STEM). The university's undergraduate enrollment is around 23,000 students, the majority of whom are white and male, countering national trends indicating higher education is becoming more racially diverse and female-oriented (*National Center for Education Statistics*, 2010). The university is situated well away from major urban population centers, and the town in which the university is hosted is small, with a total population of about 37,000 people.

This section meets in a part of campus that is intended for the university's "honor students," even though the writing course is a required course open to all first- and second-year undergraduate students. The building that houses the classroom for this section is newer than most of the buildings on campus; the university's main science complex, in contrast, is older and has smaller classrooms with fewer accommodations for technology such as projectors and overhead document cameras. This classroom is large enough to seat its 15 students and provide enough room at the front for the teacher to work. There are whiteboards on two walls, with a screen for the projector positioned over one of the whiteboards to indicate the front of the classroom where the teacher is usually situated. Long tables form a large rectangle with seats positioned along the outside of the rectangle so that class participants can see each other.

Class sessions are 50 minutes long and generally focus on a particular element of academic writing, such as primary sources or formatting of bibliographies, rather than on providing silent time for writing, which is done at home and posted on the course's online portal. When homework writing is presented in class, it is typically shared amongst peers for feedback. Any writing actually done in class is typically considered "generative writing" as a means to help students plan their essays or reflections at home.

Participants

The teacher, given the pseudonym Emily, is a doctoral student in the university's education program and a first-year instructor in the writing program. Emily has work experience in higher education contexts, but this is her first teaching experience in a public university. In interviews conducted with her, she says that she applies a lot of what she learned from her doctoral studies into teaching her writing class. The narratives used in this paper relate to experiences in her second semester of teaching.

There are 15 students in the section of the writing course observed for data collection. Of those that gave informed consent, three students are perceived African-American females, one is a female international student whose native language is Mandarin Chinese, and the rest are a mix of perceived white male and female students. All students in the class are first- and second-year undergraduate students in various majors.

Excerpts of interviews with two students are used in this paper for analysis and discussion. Jeremy is an English and theater major who transferred from a small fine arts school. As perceived in participant observations, he participates in class more often than do most of his classmates. He describes himself as a good writer and doesn't consider the writing class a significant challenge. However, he doesn't consider himself a "good test-taker," and not getting a good score on the necessary Advanced Placement exam is what required him to enroll in the writing course. He participates in local theater productions as an actor and aspires to write a novel. Perceived by the researcher as white, Jeremy indicated in his interview that he was a quarter-Native American. The

discussions that revolved around essays about multiculturalism especially resonated with him as a result.

Lonnie is a first-year history major whose goal is to become a history teacher. He was born in Mexico to a Mexican mother and a white, American father. He identifies more with his white heritage and noted in his interview that his ancestors fought in the American Revolution and the Korean War, facts which serve as his motivation for studying history. He was fluent in Spanish until his family moved to the United States, when he believes he began to lose fluency in Spanish due to lack of exposure in the language. He was homeschooled until entering college, making university his first experience in a formal education context. Unlike Jeremy, who is local to the area, Lonnie struggled with his classes in the first semester because of adjustment issues relating to homesickness and unfamiliarity with formal education.

These students and their teacher became the focus of this paper because of their frequent interaction in class. The content of the interviews with these two students were based on these interactions.

Researcher lens

The researcher in this study is a second-generation Asian-American and doctoral student. His parents come from the Philippines and he was born and raised in the United States, knowing only standard American English until he studied Japanese in college. While he was pressed by his parents to do well in school – he earned an honorable mention in a creative writing contest in high school – he was a marginal student in college who struggled to adapt to discussion-based learning employed in recitation sessions for his science and political science courses. Part of the personal rationale for this study was to explore how pedagogies that aimed to facilitate academic socialization in higher education could produce benefits to learners when traditional, more passive models of learning were just as present in the classroom.

After graduating college, the researcher became a teacher at various language schools in Japan for three out of the first five years of his professional career (he returned to the United States for two years to earn a graduate degree in Teaching English to Speakers of Other Languages), after which he became a high school and university teacher of English as a foreign language for nearly eight years. Throughout his professional development, he became aware of pedagogies that focused more on active learning as a means to elicit spoken production from language learners in a communicative environment. He then adapted scholarly work and professional advice on the subject to his own pedagogy to encourage language learners to speak and to overcome anxiety to communicate in a foreign language. He returned to the United States to pursue a doctoral degree in education with a research emphasis on academic socialization and language socialization in higher education contexts. The experience gained in promoting active learning in his own classroom makes up part of the lens that informs the analysis of the research in this paper.

Methodology

After receiving permission from the course instructor and her department, the researcher conducted weekly observations of the class section beginning in the second week of the university's spring semester. During observations, the researcher would sit among the students and monitor whole-group and small-group activities. When asked, the researcher participated in activities and answered questions that students may have. Field notes about observations made during class were recorded in a written journal. Classroom observations were about 50 minutes long.

In addition, semi-structured interviews with the course instructor and select students were conducted outside of class sessions and recorded using a digital voice recorder. Interviews with students were typically 30 minutes long, while interviews with the teacher lasted at least 60 minutes. Field notes and transcribed interviews were coded and analyzed for themes and narratives that might prove useful for discussion about active learning and teacher facilitation in the classroom.

Data analysis

Preliminary composition of field notes and structuring of interviews were conducted with an approach toward examining academic socialization through a framework adapted from theory on language socialization (e.g., Ochs & Schieffelin, 2001).

Ochs' (2004) framework can be adapted to analyze what takes place in a college classroom. The following data exemplar, represented in Table 1, looks at a field note from one of the class sessions and defines elements that fit into the four elements of the established framework.

Table 1. Analysis of data exemplar through Ochs' (2004) framework.

Teacher groups students together for a group activity. Each group has a different piece of information about Hurricane Maria, provided by a primary source. Questions to answer about the information are shown on the PowerPoint slide. (PO – 03/22/2018)

categories	• primary sources
	• type of information (e.g., newspaper article, blog post)
rules	• answer the questions
	• present the information according to the questions asked
expectations	• present to the whole class in a manner that they can understand
strategies	• determine consensus on answers to questions
	• negotiate how to present
	• decide who presents

In the class activity highlighted in Table 1, small groups of students discuss a piece of information taken from news sources. The terminal goals for this activity are both explicit through rules (i.e., answer the questions and present the information) and implicit through expectations (i.e., present to the class in a meaningful way). Less stated and less prescribed by the teacher is how students are to most successfully participate in this activity; those elements are outlined by the researcher's understanding of strategies at the time of the activity. In this case, the researcher believes that students need to come to a consensus on the answers to the questions on the PowerPoint slide, decide who presents the information, and how that information is presented to the whole class. Through this analytical framework, the researcher can examine discourse practices, behaviors, and dispositions and frame them in explicit terms conducive to discussion of classroom interaction.

Field notes and interview excerpts are analyzed to form narratives relevant to the themes presented in the findings section of this paper. They are analyzed through the framework provided above and rendered into narrative form for the purposes of this paper. The narratives presented in this study are intended to express how the teacher and her students negotiate the elements of socialization in their writing class. With respect to the teacher specifically, the narratives also focus on what the teacher does to facilitate understanding and encourage participation from her students.

Findings

Three major themes were seen by the researcher as evidenced in the data collected for this study. These themes speak to what the teacher does, but excerpts of interviews also touch on perceived effects or influence. During data analysis, it was apparent to the researcher that the teacher sought to achieve a more equal power dynamic with students and encourage their input by explicitly validating their participation, validating their knowledge, and relinquishing their own role as an authority figure.

Validating student participation

> *Teacher asks about example outline of essay shown on PowerPoint slide. One student answers often. After four answers, teacher responds, "Yeah, you're on it today!" (PO – 04/04/2018)*

During sessions, the classroom is typically quiet in the absence of group or pair work as students' attention is fixed on the teacher or disengaged from the class work altogether. In this instance, Emily is reviewing a graphic organizer of an essay outline displayed on the projector. The outline refers to an essay printed in one of the course textbooks which students were supposed to have read prior to the class session. Emily asks about the purpose of each paragraph detailed in the outline, waiting for students to raise their hands and answer. Only a handful of students volunteer answers over the course of the semester, while most students in class remain silent unless directly called on. Steve, one of the white male students, raises his hand to answer one of Emily's questions after a long silence. Emily is satisfied with his answer, as well as the three answers that he gives in quick succession when no one else chimes in.

  In the view of the researcher, this positive reinforcement is intended not only to give praise for the perceived "correct" answer, but for the willingness to participate as well. Especially when the class in this whole group discussion is fairly quiet, the researcher sees that it is important to provide positive reinforcement to those students who do participate. The act of praise in this case is meant to encourage further participation and engagement with the knowledge being presented.

  Of course, student participation must be seen by the students as rewarding and having a tangible benefit Emily references the same instance above in her interview and notes the progress Steve has made as a result of such participation.

> Interviewer: Now, you said that, um, it sounds like you've seen a turning point with this student? That they are able to write now?
>
> Emily: He had a big comeback, he had a period of time where, um, so, unit three, they did research papers. And he, like, I don't know, he kept sleeping through his alarms and, like, not coming to class, and not doing assignments. He was just out to lunch for, like, half of unit three. But in the end, he made this big comeback. I don't know what it was. Actually, you were in the class that day. He was sitting there, we were talking about reverse outlining, and he kept raising his hand, like, all the time. And I was like, "You're on it today!" And he wrote in his reflection, he was like […] so, they do tell me what they're thinking about in their reflections. This told me when I read his thing that something happened a little bit. "I feel like I learned a lot in this unit, and I realize that I am a better writer than I think I am. At the start of the unit, I

would put off a lot of work, because I would think it would be very troublesome, and I wouldn't be able to think of stuff to write. As we progress with the unit, I saw my writing improve more and more. It wasn't just writing a boring research aspect paper, but it was a topic I had some interest about. When I eventually decided to stop putting stuff off and sit down and write, the words would just start flowing in my head. I learned from my peers different techniques to think of ideas and different approaches to writing. Also, I started participating more in class, and I realized that the answers that I've been thinking of but keeping in my head were the right answers, and that I had no need to fear of being wrong. I feel, like, in the end, I develop more as a writer, and I gain more confidence when people were praising my work."

Interviewer: Wow. How do you feel about that?

Emily: I feel good. I mean, I feel, like, happy. I mean, he may just be, like, "tell me what I want to hear." But, um, I feel good that, because I think when I read that, it doesn't sound like someone who says, "I hate writing." It sounds like he actually feels more motivated and he developed more of an interest in doing it. I think, for a student like him, he needs to build confidence, that he can do it, and his paper was actually pretty good.

Student participation in the writing process is validated just as is student participation in the class. Composition writing, at least in this class, focuses on the process as much as, if not more than, the product itself, requiring creation and revision through feedback. In the following interview excerpt, Jeremy describes the writing process for a previous assignment, and notes how feedback from Emily and subsequent engagement in the revision of his essays have served a purpose.

Interviewer: Is there anything new that you have learned in this course so far?

Jeremy: Yeah, you know, I really like how she sets it up where there are multiple drafts and it's kind of like, the writing is a work in progress until you get to that final draft because most of the English classes I've had, you just hand in the final draft. You have to figure out all the rough drafts and stuff for yourself beforehand. So, I haven't gotten nearly as much feedback as I've gotten in this class. I feel like there's been a lot of feedback, which I like. I appreciate that. So, it's different, but it's useful for me. So, I like that. Yeah, that's something new.

Interviewer: Was there any part of the course, maybe big or small that you thought, "Oh, this is more challenging than I thought it would be?"

Jeremy: Uh…more challenging? I guess it would be…the homework itself isn't very difficult, but I guess the frequency of the homework, I wasn't expecting as much. Just 'cause the majority of English classes I've taken, it's a lot of reading and there are, like two or three essays, but they're like big essays, they're very important for your grade. Um, just the frequency of homework, the fact that there's a lot of it, it's more challenging. You know, just keeping on track.

Lonnie corroborates the expectations set by Emily as described by Jeremy. Students *must* participate in the course in an active manner and are rewarded in a way that helps them further in the semester, as Lonnie describes in the following interview excerpt.

Interviewer: The writing class, is it academically challenging?

Lonnie: Yeah, I would say it's not…hard in terms of studying, but it is hard in terms of the amount of work that you do. Uh, there is a lot of stuff you need to keep track of, and if you miss any of the assignments, you are already at a disadvantage. Not necessarily in terms of your grade, but just in terms of writing your paper, there's a lot of stuff in the assignments that helps you write a better paper for what you want to go for, that you may not have ever considered or ever heard of before.

Interviewer: Okay, and…do you feel like you would have been able to write an essay for your history class without taking the [writing] course?

Lonnie: Oh, yeah, definitely. Um, it wouldn't have been as good, though. I've had a lot of, um, experience writing history essays before, my AP classes and everything. But I think that taking this class has given me, like, a fresh perspective on how you should go about looking at an essay in the first place. Then, not even to get started writing it. And so definitely I could have done history essays, but I don't think I would have done them as well.

The writing course as perceived and depicted by Lonnie and Jeremy emphasizes rigor and reflection throughout the writing process. As Lonnie indicated in an excerpt presented earlier in this paper, he is capable of sitting down and writing an essay in one short period of time. The product from such a rushed effort would be, in his view, acceptable. However, being able to write an essay in a comprehensive process that involves drafting and outlining, from both students' perspectives, is seen as beneficial. The full process of writing, in the expectations that Emily sets, does require participation; one simply cannot, for example, read a book or passively absorb information about writing and become successful in producing an essay. As Lonnie says, the writing that he produces without rigor would be acceptable, but would not be as good as an essay that is produced through the participation that Emily recommends.

The feedback process, while aimed at producing better writing, or at least more disciplined writing, also has a more profound effect on Emily's students, as perceived by Jeremy in the following interview excerpt.

Interviewer: Have you ever, um, yeah, what's your general opinion about this professor?

Jeremy: I really like her. I think that she cares about students doing well. She frequently gives feedback. She's very good at giving you different chances. I feel like I'm doing well, but it seems to me, like, if a student wasn't doing well, she would do everything to help. It seems to me like she wants everybody to succeed, and she's passionate about what she's teaching. And she's friendly.

Interviewer: Um, as in giving different chances and wanting to help students succeed, can you give an example?

Jeremy: Sure. One of the students, I can't remember her name. She's like an international student, she doesn't speak English very well. I get the sense that she, you know, I've seen her talking with her and kinda giving her feedback and wanting her to raise her hand and have a chance to speak up. So, like, giving students who may be, like, having a hard time, whether it's based on the language barriers, which is not their fault. Um, a chance to speak up and be more integrated in the class I think is really admirable and I like that she does that.

The ritual of feedback in the writing class, whether feedback is given by Emily or by classmates, seems to create the expectation of shared interest and support within the classroom and among the students. Through this excerpt, Jeremy depicts a culture fostered by Emily of one that is welcoming of all students. Class rituals such as feedback are mentioned in the same conversation as examples that highlight empathetic practices that take into consideration the different characteristics of the learners in the class.

Literature on culturally responsive pedagogy from scholars such as Warren (2018) has drawn connections between a teacher's capacity for empathy and their capacity to learn about students as a result. However, this anecdote provided by Jeremy presents teacher empathy as an exemplar for classroom practices that are intended to have a material benefit (in this case, improvement of students' writing practices) but also have the effect a fostering a more open and respectful environment within the classroom.

Validating student knowledge

> *Teacher reviews insider/outsider concept for ethnography essays. Teacher elicits topics that all students have chosen and asks questions about each topic. "So what do you want people to know about vegans?" (PO – 04/11/2018)*

Emily chose ethnographies as the third assignment of four for the writing course. In this unit, students are expected to conduct research on a community of their choice, in consultation with Emily, and write an essay providing insight into that community. The challenge of ethnography, Emily emphasizes, is to depict the subject in terms that outsiders would be able to understand.

Students come to class having chosen a community to write about. She goes around the class and asks each student what they have chosen. Each topic is written on the board. "Fighting game community," "student athletes," "independent artists," "vegans," and "Chinese students" are among some of the topics that the students call out.

"What do you want people to know?" is the sort of question that Emily asks to some of her students about their chosen communities. Sometimes, it is intended to clarify the community, as was the case with "Chinese students," intended to mean Chinese international students at the university. Other times, it is meant to preface what each student is later asked to do, which is to pair up with another student unfamiliar with the topic and discuss. Emily asks "What do you want people to know

about vegans?" to elicit a brief answer to highlight the requirement of explaining unfamiliar concepts to outsiders.

In any case, this class activity is centered on the students' knowledge, deemphasizing what the teacher knows and what the teacher expects to read in students' essays. Other examples presented thus far have highlighted how Emily attempts to focus the class on the input that her students contribute. Unlike rituals such as peer feedback, however, this activity is a teacher-to-student interaction. For this researcher, this serves as a means to publicly validate everyone's contribution in front of everyone else. Just as struggles in whole class activities magnify students' anxiety and threaten their willingness to participate (as seemed to be the case presented earlier in this paper when Lonnie missed a question), activities such as these provide public positive reinforcement of students' ideas.

As much as these sorts of activities encourage students to participate and contribute their insights, they produce a community-building effect as discussed previously and in discussion of the next field note.

> *Teacher pairs students so that each student is unfamiliar with their partner's chosen community. Students reseat and share what they know about their chosen communities with each other. Richard talks about car culture with Allie, who talks about mixed Trinidadian culture. Allie appears genuinely interested, asking questions about people who are into car culture. Both exchange looks of interest about the other's chosen community. (PO – 04/11/2018)*

The field notes collected for this study document many instances where students appear unengaged and remain quiet in whole group activities. In contrast, the pair work and small group activities provide opportunities away from the direct and constant monitoring of the teacher to have students interact with each other in ways that they probably would not do otherwise.

This does not mean that the fostering of communication from and among students is wholly unstructured, against which Boud et al. (2001) provide strong warnings. Emily plays a role that may not have an immediately obvious purpose to the class. Across participant observations, students are seen to sit in the same seats, sitting next to the same students in class each session. At times, Emily counts off students into groups to reseat them. Absent the reseating, students have little other opportunity in class to talk with other students besides those next to them. In the above field note, Allie and Richard sit at opposite ends of the classroom, as Allie, a perceived African-American, invariably enters each session with her other friends. However, Emily structures the pair work activity so that students learn new insights that they would be unlikely to learn outside of the activity.

The next excerpt of Jeremy's interview highlights the effects of community-building activities, which appear to transcend the immediate, material benefits of this knowledge sharing in class.

> Interviewer: When you say discussion-style activities or discussion in class as works for you, are we talking about in the whole class or in small groups?
>
> Jeremy: I think both. I really like whole class discussions because I like hearing what other people have to say, so I think there's some fascinating things that people can bring to the table, their own views on certain topics. Um, and I like having, I think it makes the class closer when you have, you know, people that can come out and, kind of, give in their two cents. Um, so, yeah, I've been always into that. And then

> also like the small group discussions, um, I learn a lot about other people, other students, kind of where they're coming from and how their thinking may not align with my own.

The small group and pair work activities that Emily has structured appear to tap an intrinsic motivation in at least some of the students, who are less likely to engage (or are at least less interested) in the whole group activities such as presentations. Later in the interview, Jeremy indicates how the presentations to the class, which are prepared in small group discussions, appear to him as simply regurgitating key points, reducing them to a mere recitation that does not interest him. Whether it is lack of interest or, perhaps for other students, anxiety to contribute in front of the entire class, activities that involve groups of only a few students each can provide a safe space to share ideas, while the reseating of students by the teacher can elicit the opportunity to learn unfamiliar ideas.

This exchange of ideas, however, is not possible without the validation of the students' knowledge. Traditional models of education, as typified by the lecture, favor the expertise of the teacher much more than they do any expertise provided by their learners. In this classroom, however, Emily appears to emphasize what her students have to say by giving those ideas a platform in the various classroom activities she has designed. This speaks to another theme that relates to differences in power and authority, as discussed in the next section.

Relinquishing role of authority

The students in the writing class call their teacher by her first name. In contrast, when another course is mentioned, particularly a required lecture course common to the STEM majors such as Organic Chemistry or Introductory Biology, students refer to the teacher by their family name, complete with a title such as Mr., Ms., or even Dr. The decisions made regarding how to refer to a person draw on a speaker's assumptions of power and status. When the students refer to Emily by her given name, a more symmetrical connection between teacher and student is assumed.

Educators do not enter the classroom on equal footing with those of their students. Teachers are seen as authoritative sources of knowledge necessary for students to complete coursework. Parallel to this, they award grades for good performance and can fail a student who does not master the presented content knowledge or is otherwise noncompliant or cooperative during class. Quite simply, a student can face difficulties if their teacher does not approve of their performance. A teacher in this case can be the gatekeeper of expertise status as well as academic success.

This perceived circumstance can create a large divide between students and their teacher. When Lonnie remarks about getting "roasted" a second time after giving an answer that wasn't seen as accurate, he appears to be reacting to the exercise of authority, however gentle, by the teacher. This suggests the likelihood that interactions in the class will involve asymmetrical power dynamics especially if the teacher plays a role. Intervention through pedagogy, as discussed in previous sections, can influence how the teacher can mitigate her role as an authority figure in order to validate what students bring to the classroom.

> *Brief discussion between activities centers around meaning of "savage" in casual conversation. Teacher asks students what it means. (PO – 02/21/2018)*

Here, Emily takes a moment created by a natural break between class activities to refer to a previous discussion of an essay about World Englishes and how different communities create different varieties of English which may, in turn, seem foreign to other communities.

In the instance described by the above field note, Emily attempts to elicit some conversation into the meaning of "savage," a colloquialism among her younger students, as an example of differing varieties of English, as used by students in her class in the previous semester. In doing so, her unfamiliarity with the word in such context indicates how she and her generation doesn't use the word in the sense that her students more likely do.

For one, a lack of expertise dispels any potential notion that the teacher knows all. At first glance, this is not significant, since she is the perceived expert on writing in the class, not the expert on varieties of English. However, as she asks her students what the word is intended to mean, she is validating what knowledge might be important and, at the same time, validating any knowledge her students have on the subject.

For reference's sake, the Urban Dictionary entries for "savage" suggest it is a synonym for "cool" or "badass" (n.d.), and the end result of this brief exchange involves some of Emily's students giving some example uses involving the word, clearing up some confusion, if not all. However, what is important to take away from this instance is Emily's deference to the student's expertise, decentering her authority and valuing her students' contributions.

As for knowledge that the teacher might be expected to know and convey to students, the perceived lack of knowledge on the teacher's part can provide a useful opportunity to validate students' knowledge and mitigate differences in power.

> *Teacher mentions the difference between "pp." and "p." in MLA citation. Asked why there's a difference. Says "I don't know why, there's probably a reason." (PO – 02/28/2018)*

In the session that focuses on copyediting, Emily spends a moment pointing out that "pp." in an essay's reference list is used when multiple pages are used, as opposed to the use of "p." to indicate the use of a single page. When asked the reason for the distinction by a student, Emily doesn't have a conclusive answer. Rather, she re-emphasizes what MLA style, the copyediting style that is used for essays in her class, prescribes.

The opportunity presented here is one in which Emily, who has exhibited expertise in other instances, can remove herself from the position of authority and express the same uncertainty as do her students. Rather than act an authoritative force in the classroom, Emily can be seen as existing on the same level as her students, searching for the same answer that one of her students is looking for. While she is still the gatekeeper, ultimately deciding how her students' essays will be received and what grades her students will have, her self-effacement in not knowing a particular answer serves as a substitute for validating students' knowledge, lowering her own status and mitigating the power distance between her and her students.

Reflections

This paper examines pedagogy as a means to improve learner outcomes in higher education. However, policy implications are not a central focus of this discussion, as this study is an attempt to highlight what individual practitioners in higher education can do to reach students who come from different backgrounds and circumstances. The researcher for this study recognizes that class sizes considerably larger than the fifteen students in Emily's writing course will continue to be a presence in major universities, at least in the foreseeable future. As such, there are concerns of the practicality in hiring and educating more instructors to ultimately reduce class sizes at the college level. Despite this, research (e.g., Goldacre et al., 2013) applying at least some of the principles of active learning and

rapport building to large university classrooms holds some promise for lecturers connecting with greater numbers of students. This paper is thus an attempt to depict the possibility that any classroom can be collaborative and welcoming of the insight that learners bring to higher education.

Implications

Teachers in higher education contexts are likely to encounter students with interests and goals that are disparate from those of their peers. Jeremy, the theater major and actor, sees different goals for his academic writing than does Lonnie, a history major and aspiring future teacher. Connecting with both students, and with the other students in class, is a task that a method of instruction that treats all learners the same may not be able to accomplish. Practitioners who are able to practically find time to connect with their students on an individual basis (primarily because of manageable class sizes), should consider the following questions to inform their pedagogy:

- How much time and effort in the course is devoted to the insights that students bring to the classroom?
- How is student knowledge elicited in class in a way that is non-threatening to learners?
- To what extent is peer interaction with respect to class material encouraged and structured?
- To what extent is the teacher's authority and expertise emphasized or de-emphasized to their students?

As for theory on academic socialization, the framework developed for this study might have benefited from a deeper exploration of cause and effect in academic settings. While the framework provided by Ochs (2004) provides some useful definition to examining rituals and practices within the classroom, it appears to miss elements of analysis that would have been able to explore why students and their teacher make the decisions that they do. In Ochs' softball example, the child struggled with understanding the strategies employed while batting and running the bases. In a stakes-free situation, however, behavior may have differed from that which may exist in an academic situation with larger implications. Ethnographers researching academic environments may want to expand their theoretical frameworks by considering the following questions:

- What are the stakes involved in participating in the academic community being observed?
- What motivations (intrinsic or extrinsic) do novices such as in the academic community being observed have when participating?
- What motivations do experts such as practitioners have when fostering participation by novices?

Limitations and future directions

Data collection was limited to participant observations and interviews; replication of this research should be complemented by document collection (e.g., collection of essays and other written work) and audio recordings of class sessions for transcription and discourse analysis. The latter aspect of data collection would have served a purpose in tying pedagogical practices to discourse practices on the part of the teacher.

Research in teaching pedagogies could have also be accompanied by direct observation of effectiveness in addition to perceptions of effectiveness. While the two interview subjects express

general satisfaction with the course and perceive gains in their own writing ability and organization in a way that is attributable to the course or the teacher, the research into teaching methods in higher education contexts would benefit from an empirical examination of effectiveness on academic performance.

Finally, while this paper is framed as a discussion of alternatives to traditional teaching methods in higher education in the United States, the comparisons drawn between lecture and more dynamic pedagogies were only meant to serve as a framing of active learning as an innovation not yet fully embraced by university educators. Contrastive studies that explore the differences in approach to pedagogy as well as effectiveness are welcome in exploring larger questions regarding what methods of instruction are most appropriate in the various facets of university education.

Conclusion

In summarizing the findings discussed in this paper, this study of one university-level writing class identifies a teacher's strategies for establishing rapport, particularly through valuing the participation and knowledge that students contribute to the class while also de-emphasizing the expertise she brings as an authority figure. More important than the discrete acts or utterances produced by the teacher is the general philosophy guiding the building of rapport within the classroom, to which students appear to respond favorably and which has generated perceptions of pathways to student success.

There is no shortage of literature regarding learner-centeredness, active learning, or culturally responsive pedagogy. However, given the stakes of improving learner outcomes in higher education in the United States, researchers would benefit education by more closely examining alternatives to the status quo in pedagogies in the university classroom. Discussion of reforms intended to ensure more equitable outcomes in higher education are always incomplete without reforms to curriculum and campus climate in a manner that ensures a greater sense of inclusivity accommodating all students. Moreover, the ethnography presented in this paper, at least in the view of the researcher, did not encounter instances of overt discrimination or marginalization of underrepresented students that culturally responsive pedagogies would be suited to address. Therefore, it would be presumptuous to assume that the pedagogical implications presented here can serve as a panacea to all the ills that college classrooms in the United States might face.

The only assertion that can be reasonably made by this research is that teacher practices that establish rapport, center education around the needs and knowledge of learners, and equalizes power dynamics between teacher and student are the precursor to pedagogies that respect and negotiate differences among all learners. At minimum, the findings presented through this research should be part of a larger discussion about the nature of the university classroom and what teaching and learning can look like in order to produce the most effective learner outcomes in higher education.

References

Arghode, V., Wang, J., & Lathan, A. (2017). Exploring professors' engaging instructional practices: A collective case study. *Journal of the Scholarship of Teaching and Learning, 17*(4), 126-149.

Boud, D., Cohen, R., & Sampson, J. (2001). *Peer learning in higher education: Learning from & with each other*. London: Kogan Page Limited.

Chickering, A. W. & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *American Association for Higher Education*, *39*(7), 3-7.

Colbert, P. J. (2010). Developing a culturally responsive classroom collaborative of faculty, students, and institution. *Contemporary Issues in Education Research*, *3*(9), 17-26.

Estepp, C. M. & Roberts, T. G. (2015). Teacher immediacy and professor/student rapport as predictors of motivation and engagement. *NACTA Journal*, *59*(2), 155-163.

Goldacre, L., Bolt, S., & Lambiris, M. (2013). Designed for learning: A case study in rethinking teaching and learning for a large first year class. *Journal of the Scholarship of Teaching and Learning*, *13*(1), 24-44.

Grauer, G. F., Forrester, S. D., Shuman, C., & Sanderson, M. W. (2008). Comparison of student performance after lecture-based and case-based/problem-based teaching in a large group. *Journal of Veterinary Medical Education*, *35*(2), 310-317.

Lee, C. (2015). More than just language advising: rapport in university English writing consultations and implications for tutor training. *Language & Education: An International Journal*, *29*(5), 430-452.

Musto, P. (2017). Low-income students see low graduation rates. *Voice of America*, retrieved from https://www.voanews.com/student-union/low-income-students-see-low-graduation-rates

Ochs, E. (2004). Becoming a speaker of culture. In C. Kramsch (Ed.), *Language acquisition and language socialization: ecological perspectives* (pp. 99-120). London: Bloomsbury Publishing.

Ochs, E. & Schieffelin, B. B. (2011). The theory of language socialization. In A. Duranti, E. Ochs, & B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 1-21). Malden, MA: Wiley-Blackwell.

Reimschisel, T., Herring, A. L., Huang, J., & Minor, T. J. (2017). A systematic review of the published literature on team-based learning in health professions education. *Medical Teacher*, *39*(12), 1227-1237.

Roithmayr, D. (2014). *Reproducing racism: How everyday choices lock in white advantage.* New York: New York University Press.

Savage. (n.d.). *Urban Dictionary*, retrieved from https://www.urbandictionary.com/define.php?term=savage

Status and trends in the education of racial and ethnic minorities. (2010). *National Center for Education Statistics*, retrieved from https://nces.ed.gov/pubs2010/2010015.pdf

Taglieri, C., Schnee, D., Camiel, L. D., Zaiken, K., Mistry, A., Nigro, S., Tataronis, G., Patel, D., Jacobson, S., & Goldman, J. (2017). Comparison of long-term knowledge retention in lecture-based versus flipped team-based learning course delivery. *Currents in Pharmacy Teaching and Learning*, *9*, 391-397.

Tate, E. (2017). Graduation rates and race. *Inside Higher Ed*, retrieved from https://www.insidehighered.com/news/2017/04/26/college-completion-rates-vary-race-and-ethnicity-report-finds

Thistlethwaite, J. E., Davies, D., Ekeocha, S., Kidd, J. M., MacDougall, C., Matthews, P., Purkis, J., & Clay, D. (2012). The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. *Medical Teacher*, *34*(6), e421-e444.

Thrall, G. C., Coverdale, J. H., Benjamin, S., Wiggins, A., Lane, C. J., & Pato, M. T. (2016). A randomized controlled trial of team-based learning versus lectures with break-out groups on knowledge retention. *Academic Psychiatry*, *40*(5), 755-760.

Verner, C. & Dickinson, G. (1967). The lecture: An analysis and review of research. *Adult Education*, *17*, 85-100.

Warren, C. A. (2018). Empathy, teacher dispositions, and preparation for culturally responsive pedagogy. *Journal of Teacher Education*, *69*(2), 169-183.

Wlodkowski, R. J. & Ginsberg, M. J. (1995). A framework for culturally responsive teaching. *Educational Leadership*, *53*(1), 17-21.

Small Group Learning is Associated with Reduced Salivary Cortisol and Testosterone in Undergraduate Students

Kristin Snopkowski
Boise State University
kristinsnopkowski@boisestate.edu

Kathryn Demps
Boise State University

Shane Scaggs
The Ohio State University

Ross Griffiths
Boise State University

Karen S. Fulk
Boise State University

Scott May
Boise State University

Kimberly Neagle
Boise State University

Kayla Downs
Boise State University

Michaela Eugster
Boise State University

Tessa Amend
Boise State University

Abstract: Small group learning activities have been shown to improve student academic performance and educational outcomes. Yet, we have an imperfect understanding of the mechanisms by which this occurs. Group learning may mediate student stress by placing learning in a context where students have both social support and greater control over their learning. We hypothesize that one of the methods by which small group activities improve learning is by mitigating student stress. To test this, we collected physiological measures of stress and self-reported perceived stress from 26 students in two undergraduate classes. Salivary cortisol and testosterone were measured within students across five contexts: a) pre-instructional baseline, b) following a traditional lecture, c) after participating in a structured small group learning activity, d) following completion of multiple choice, and e) essay sections of an exam. Results indicate students have lower salivary cortisol after small group learning activities, as compared to traditional lectures. Further, there is no evidence of a relationship between physiological measures of stress and self-reported perceived stress levels. We discuss how structured small group activities may be beneficial for reducing stress and improving student-learning outcomes.

Keywords: college students, stress, endocrinology, instructional techniques

Previous research has shown that active student learning activities, such as those that occur in small groups, improve student performance and engagement as compared with traditional instructional lecture (Bradford, Mowder, & Bohte, 2016; Byun, 2014; Coakley & Sousa, 2013; Freeman et al., 2014; Simonson & Shadle, 2013; Swap & Walter, 2015; J. D. Walker, Cotner, Baepler, & Decker, 2008; L. Walker & Warfa, 2017). However, little is known about the mechanism by which this occurs. Identifying these mechanisms may lead to improved learning outcomes both within the classroom and beyond.

We hypothesize that one of the methods by which small group activities improve learning is by mitigating student stress. To compare the stressfulness of group learning to traditional lecture we use measures of salivary cortisol as a physiological indicator of stress in undergraduate courses. We also examined students' physiological stress responses during multiple-choice and short-answer essay sections of an exam. Finally, we compared salivary cortisol to self-reported levels of stress. Since most research on student stress and learning rely on student self-reports, it is important to document how closely these correlate with physiological measures of stress.

Stress, Cortisol, and Learning

Although cortisol has often been called the 'stress hormone', a more accurate statement would describe it as the 'arousal hormone' (Hoyt, Zeiders, Ehrlich, & Adam, 2016). Cortisol is a glucocorticoid hormone that is released when the hypothalamic-pituitary-adrenal (HPA) axis is activated. Circulating cortisol levels increase in response to physical and psychological activation; thus, it is used as a biomarker of a stress response (Kirschbaum, Pirke, & Hellhammer, 1993; Lighthall, Gorlick, Schoeke, Frank, & Mather, 2013; McEwen, 1998; Stephens, Mahon, McCaul, & Wand, 2016). While other measures of stress exist, including serum cortisol, galvanic skin response, heart rate, and blood pressure (Campbell & Ehlert, 2012; Villanueva, Valladares, & Goodridge, 2016), these other measures typically are utilized within laboratory settings due to the need for aseptic techniques, machines, or continuous sensor leads connected to individual participants throughout the evaluation process. In contrast, the use of salivary cortisol readily lends itself to the formal classroom setting where less-invasive and least restrictive environmental conditions are desired to enable the simultaneous evaluation of many students.

Stress has been demonstrated to have positive and negative effects on learning and memory (Sapolsky, 2004). While too little stress can elicit boredom (Merrifield & Danckert, 2014), prior research indicates that a sufficient level of stress enhances learning, but only for positive learning outcomes (Lighthall, Gorlick, Schoeke, Frank, & Mather, 2013), such as those associated with instructional reinforcement of correct responses. However, excessive and repeated stress can reduce cognitive function, increase the risk of cardiovascular disease, and decrease immune function (Lee et al., 2007; McEwen, 1998; Robinson, Sünram-Lea, Leach, & Owen-Lynch, 2008). Evidence has shown that inducing stress results in increased cortisol levels and impaired memory (Kirschbaum, Wolf, May, Wippich, & Hellhammer, 1996). In overly stressful learning contexts, learners may be activating the sympathetic nervous system, often called the "fight-or-flight" response, increasing circulating cortisol, and impairing their ability to retain course material.

Studies using salivary cortisol as a biomarker of stress have found that exams can be used as a naturalistic experiment to examine the effects of academic context on stress. For example, students had increased levels of salivary cortisol immediately prior to oral exams (Lacey et al., 2000; Schoofs, Hartmann, & Wolf, 2008; Singh et al., 2012). Similar results have been found for undergraduates facing

oral presentations (Merz & Wolf, 2015). In fact, asking people to do math problems in front of an audience is part of the Trier Social Stress Test (TSST), a reliable and validated method to induce stress in laboratory settings (Allen et al., 2017; Kirschbaum et al., 1993). While oral exams consistently increase salivary cortisol, written exams have more variable effects. For instance, Austrian high school students had varied reactions to written exams (some increased, some decreased, and some showed no change in salivary cortisol levels) (Martinek, Oberascher-Holzinger, Weishuhn, Klimesch, & Kerschbaum, 2003), while British university students had a significant reduction in salivary cortisol levels during exam weeks as compared to non-exam times (Vedhara, Hyde, Gilchrist, Tytherleigh, & Plummer, 2000), and German undergraduates saw elevated cortisol concentrations at the start of a written examination compared to a control day – if the control day was after the exam, but not if it was before (Preuß, Schoofs, Schlotz, & Wolf, 2010).

Many studies have focused on the physiological stress induced by examinations, but less is known about how student stress, both physiological and self-reported, varies across learning contexts. One study examined medical students in a problem-based learning curriculum and found that students reported via a questionnaire that the group-learning environment caused little stress (contrary to many other aspects of their program) (Moffat, McConnachie, Ross, & Morrison, 2004). No prior research (that we are aware of) has examined the effect of classroom learning context on the physiological stress response as measured by salivary cortisol.

Finally, the relationship between self-reported perceived stress and physiological markers of stress is unclear. In some studies, salivary cortisol is unrelated to perceived stress. For instance, studies comparing undergraduates before and during exam weeks found no association between salivary cortisol and self-reported stress (Murphy, Denis, Ward, & Tartar, 2010; Weekes et al., 2006). In others, researchers found a positive relationship between self-reported stress and biomarkers of stress (Ng, Koh, & Chia, 2003). Finally, salivary cortisol can also be negatively correlated with self-reported stress. Acute stress can lead to the release of endorphins, which reduces the perception of pain (for example, the runner's high) (Sapolsky, 2004). Under these conditions, stress actually results in the perception of positive feelings. Given the complexity of the HPA axis and its reaction to stressors and other neurobiological events and the complexity of assessing self-reported stress including types of measures, previous exposure to the stressor, and an individual's coping mechanism, it may be unsurprising that salivary cortisol is not correlated with self-reported perceived stress in many studies (Hellhammer, Wüst, & Kudielka, 2009). If self-reports of perceived stress are not associated with physiological stress (or associated in a convoluted way), this may alter how we both interpret self-reports of "stress" and understand how it relates to learning. Differing pedagogical approaches to higher education create different environmental contexts that may impact the stress experienced by students, thereby impacting memory and learning. Importantly, whether these modalities directly affect student hormonal stress levels requires investigation beyond subjective student perception.

Salivary Testosterone

We also examined salivary testosterone across these learning and exam-taking contexts. Ellison and Gray (2009) determined that cortisol and testosterone are implicated in the physiological stress response and noted that these hormones are tied to both individual and group learning. Previous research has found that acute stress is associated with increased testosterone (within individuals) and positive associations occur between salivary cortisol and testosterone, known as 'coupling' (Harden et al., 2016). Testosterone is also known to be associated with competition, where individuals who win competitions exhibit higher testosterone levels than those who lose (Booth, Shelley, Mazur, Tharp, & Kittok, 1989). In some circumstances, rises in testosterone occur in anticipation of competition (Mazur, Booth, & Dabbs, 1992), although patterns appear to differ for males and females (Kivlighan,

Granger, & Booth, 2005; Mazur, Susman, & Edelbrock, 1997; Taylor et al., 2000). If inter-student competition for grades or academic recognition is contributing to classroom stress, we would expect to see a signature in salivary testosterone profiles of students across contexts. Because the connection between stress, learning, and testosterone is unclear, we refrain from making predictions and only report exploratory results.

Despite the potential significance of the interaction between stress, hormones, and learning (Flegr & Priplatova, 2010; Lacey et al., 2000; Lighthall et al., 2013; Martinek, Oberascher-Holzinger, Weishuhn, Klimesch, & Kerschbaum, 2003) as evidenced by a growing body of current literature exploring the relationships between hormones and student academic performance on exams (Bardi, Koone, Mewaldt, & O'Connor, 2011; Kenwright et al., 2011; Takatsuji et al., 2008; Vedhara, Hyde, Gilchrist, Tytherleigh, & Plummer, 2000), there is a need to broaden this investigation to include the hormonal implications that different teaching modalities have on students within the formal classroom setting. In doing so, additional insight may be gained into the endocrinology and context of human learning.

Methods

We collected saliva samples to measure (salivary) cortisol and testosterone. Most studies report a relatively high correlation between serum (blood) cortisol levels and salivary cortisol, particularly among individuals with normal endocrine functioning (Hellhammer, Wüst, & Kudielka, 2009). Drawing blood may increase stress (influencing the hormones we wish to measure), so saliva collection is our preferred method to measure cortisol and testosterone levels.

Saliva samples were collected from students after five different conditions. In each case, saliva samples were collected twenty minutes after the onset of the condition (except for baseline, which was collected at the beginning of class before instruction, immediately after acquiring participant consent). These five conditions included: 1) baseline sample; 2) following ~30 minutes of traditional lecture; 3) following ~30 minutes of a small group activity; 4) following at least 20 minutes of a multiple choice examination; and 5) following at least 20 minutes of an essay examination. We modeled our small group activity after the POGIL (Process Oriented Guided Inquiry Learning) instructional technique (Moog, Creegan, Hanson, Spencer, & Straumanis, 2006), but it has not earned POGIL Project endorsement. The collection of saliva after the traditional lecture and the small group activity occurred on the same day (referred to as *instructional day*). Similarly, the collection of saliva after the multiple choice questions and the essay questions in an exam also occurred on the same day (referred to as *exam day*). The order of the instructional techniques (lecture then small group) and exam questions (multiple choice then essay) remained the same across both semesters to maintain a consistent experimental protocol. The instructional day samples were collected between two to seven days after the baseline samples were collected, and the exam day samples were collected seven to ten days following the instructional day. Each class engaged in at least two previous small group activities (with the same group members) and several traditional lectures to familiarize students to these instructional methods.

Hormones can be influenced by countless factors, including time of day (Dowd et al., 2011), yearly season (Persson et al., 2008), time since last food intake (Hansen, Garde, & Persson, 2008), amount of previous night's sleep (Leproult, Copinschi, Buxton, & Van Cauter, 1997), exercise (Hansen et al., 2008), alcohol consumption (Hansen et al., 2008), education (Dowd et al., 2011), academic performance measures (Preuß, Schoofs, Schlotz, & Wolf, 2010), biological sex (Stephens et al., 2016), prior trauma (Suzuki, Poon, Papadopoulos, Kumari, & Cleare, 2014), and medications (Hellhammer, Wüst, & Kudielka, 2009), among others. Statistically controlling for all of these factors would require a substantial sample size. Given the small sample size of our study, we opted to use the same

individuals over time to control for within-subject variation, with saliva samples taken at approximately the same time each day (since saliva samples were always taken during a class period that met the same time each week). Additionally, we recorded information on each subject regarding their wake time (all three collection days) and the time of last food intake (on instructional and exam days). All subjects had been awake for over two hours at the time of collecting their saliva sample, with the exception of one individual who woke approximately 1 hour and 40 minutes before samples were taken. This individual woke at the same time on all class days where saliva was collected, so elevated cortisol levels (due to the strong diurnal pattern of cortisol peaking 30-45 minutes after waking followed by declines for the rest of the day (Dowd et al., 2011)) would be consistent across this individual's samples. Similarly, all subjects reported not eating in the hour prior to the saliva collection (on the instructional and exam days). Other factors, such as biological sex or prior trauma would be controlled for within an individual. Our statistical analyses will examine within individual changes, as opposed to mean changes of the entire sample, since we cannot control for the variety of factors that may influence salivary cortisol levels.

Sample

Our sample included two undergraduate classes taught by the same professor that met during two consecutive semesters in 2016. The classes were an introductory biological anthropology class (Spring 2016) and an upper division anthropology class (Fall 2016). Students were asked to volunteer after the goals and motivations of the study were presented (which took fifteen minutes or less at the beginning of class). All participants provided their written consent to participate. Student identification remained confidential through the use of unique codes generated by each student. The instructor of the courses was unaware of which students chose to participate. This study was approved by the Boise State Institutional Review Board (#028-SB16-036). In total, 26 Boise State students completed the on-line questionnaire (see below) and provided saliva samples (although some of these participants were absent on either the instructional or exam day). There were three students who reported taking medication for stress, anxiety, or depression who were excluded from our analyses, resulting in a sample size of 23 students (11 females and 12 males) with a mean age of 23 years.

Perceived stress was self-reported by the participant at baseline (for both courses), during the exam day (for both courses), and during the instructional day (for the fall course only). This produced a total of 23 people who self-reported their perceived stress between one and three times, for a total of 51 observations.

Saliva Collection and Questionnaires

The passive drool method was used to collect the saliva samples (Salimetrics, 2016). This process involved having the participant hold a plastic tube up to their mouth, allowing spit to slide downward into the collection vial. After collection, samples were kept frozen until sample extraction and analysis began.

On the day the baseline sample was collected, participants completed an on-line questionnaire. The survey included questions about the students' basic demographic information. On subsequent saliva collection days, participants completed a short survey on perceived stress level, time of last meal, and wake-up time (this data was not collected on the instructional day in the spring course). Self-reported perceived stress was asked using the question "How stressed out do you feel today?" and the response was measured on an ordinal 7-point interval scale ranging from "completely relaxed" to "completely stressed".

Saliva Analyses

Salivary cortisol samples were analyzed according to the Salimetrics expanded range high sensitivity salivary cortisol enzyme immunoassay kit, while the salivary testosterone samples were analyzed according to the Salimetrics expanded range salivary testosterone enzyme immunoassay kit (Salimetrics LLC, State College, PA). These included: centrifuging samples at 3000 rpm, pipetting the samples and controls into wells and adding the enzyme conjugate. After mixing the samples on a rotator plate and incubating the samples for an hour, the plate was washed four times with the wash buffer. The substrate tetramethylbenzidine (TMB) solution was added to each well and incubated for 25 minutes. Finally, a stop solution was added and the plate was read at 450 nm. For each plate, the concentrations of the controls and saliva samples were calculated by interpolation using a 4-parameter non-linear regression curve fit (Salimetrics, 2016).

Data Analysis

Since we sampled the same individuals repeatedly, within subject comparisons of salivary cortisol and testosterone formed the basis of our statistical analyses. We conducted a one-way repeated measures ANOVA to compare the effects of instructional methods and exam type on salivary cortisol and testosterone (Field, 2013). Analyses were performed in SPSS (v. 22). For pair-wise comparisons, a Šidák correction was used to maintain the familywise error rate. For completeness, both the Šidák correction and Tukey LSD (which is equivalent to no adjustment) are reported. We also examined the relationship between self-reported perceived stress (as reported by the participant) and salivary cortisol. This was done by conducting a random effects regression model where self-reported perceived stress predicts salivary cortisol and a random effect is included to control for the repeated measures design. We also provide figures of the mean salivary cortisol and testosterone values for each condition, but these do not control for the many factors that influence salivary cortisol and testosterone. We do not use these figures as the basis of our statistical tests; they are simply depicting the descriptive statistics visually.
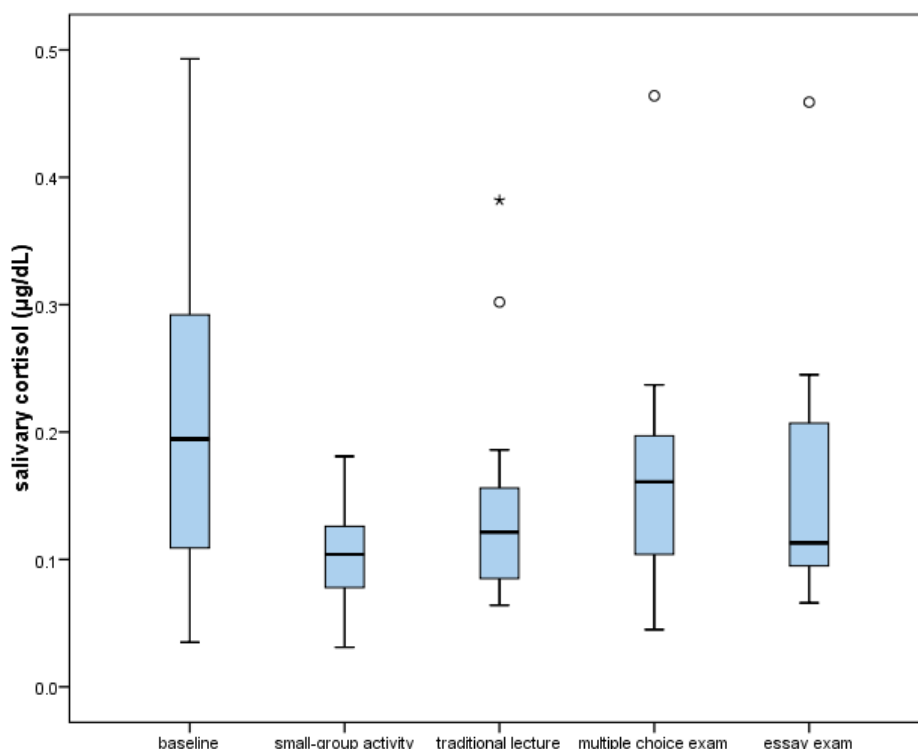
Results

Salivary Cortisol

Figure 1. Salivary cortisol (measured as μg/dL) across conditions.

Boxplots represent the upper quartile, median, and lower quartile as the top, middle, and bottom of the box. The whiskers represent the top 25% (upper whisker) and the bottom 25% (lower whisker) of data, unless one of the scores is greater than the upper quartile plus 1.5 times the inter-quartile range, which are deemed outliers and represented by an open circle (o). Values that are greater than the upper quartile plus three times the inter-quartile range are deemed extreme cases and are represented by asterisk (*). The two outliers shown for the exam day represent the same individual.

Figure 1 displays the boxplots of salivary cortisol levels after each of the five conditions. Salivary cortisol at baseline had the largest mean and standard deviation (M = 0.19 μg/dL; SD = 0.12), while salivary cortisol following a small group activity had the lowest mean value and the smallest standard deviation (M = 0.10 μg/dL; SD = 0.04). Mean salivary cortisol following traditional lecture (M = 0.14 μg/dL; SD = 0.08) was higher than mean salivary cortisol following a small group activity, but lower than mean salivary cortisol following both exam conditions: the essay portion (M = 0.15 μg/dL; SD = 0.09) and the multiple choice portion (M = 0.16 μg/dL; SD = 0.09) of an exam. This figure is suggestive of overall effects, but it is possible that they mask within-individual differences, as it is not possible to detect individual changes in summary plots.

      To examine within-individual effects, we conducted one-way repeated measures ANOVA to compare the effects of small group activity, traditional lecture, multiple choice exam, and essay exam on student's cortisol levels. Mauchly's test indicated that the assumption of sphericity (that the variances of the differences between conditions are equal) was violated ($p < 0.01$); therefore the

Greenhouse-Geisser corrected tests are reported (Field, 2013). These results showed that there are significant differences in cortisol levels across conditions ($p < 0.05$). Table 1 displays the pair-wise comparisons with and without Šidák correction of p-values. The Šidák correction was used to counteract the problem of multiple comparisons. Examining our pair-wise comparisons (to determine which groups were significantly different from each other), we found only one statistically significant pair-wise difference with a Šidák correction; salivary cortisol was significantly higher at baseline than after the small group activity ($p < 0.05$). Our evaluation of the difference between small group learning activities and traditional lecture showed that the effect was significant *only* when using Tukey LSD post-hoc test, which is equivalent to having no adjustment (see Table 1), where cortisol was lower after small group activities compared to traditional lecture (Tukey LSD post-hoc test, $p < 0.05$). The average salivary cortisol values under exam conditions (calculated as the average of multiple choice and essay cortisol values; M = 0.166 µg/dL, SD = 0.101) compared to instructional methods (average of traditional lecture and small group activity; M = 0.122 µg/dL, SD = 0.061) show that cortisol values are higher under exam conditions than after instructional methods; $t(20) = -2.065$, $p = 0.052$, which may be expected given the psychological stress that commonly accompanies testing.

Table 1: Pairwise comparisons for one-way repeated measures ANOVA for the effect of condition on salivary cortisol

Sample 1	Sample 2	Mean difference	S.E.	*p*-value Šidák correction	*p*-value Tukey LSD
Baseline	Small group activity	0.100	0.027	.016*	.002**
	Lecture	0.064	0.026	.227	.025*
	Multiple Choice	0.044	0.034	.900	.206
	Essay	0.055	0.029	.539	.074
Small group activity	Lecture	-0.036	0.014	.177	.019*
	Multiple Choice	-0.056	0.022	.200	.022*
	Essay	-0.046	0.023	.497	.066
Lecture	Multiple Choice	-0.020	0.027	.998	.474
	Essay	-0.009	0.030	1.000	.758
Multiple Choice	Essay	0.010	0.020	1.000	.614

Note: * p < 0.05, ** p < 0.01

Does salivary cortisol correlate with self-reported perceived stress?

Figure 2 displays a bar chart of mean salivary cortisol for each value of self-reported perceived stress. If self-reported perceived stress is associated with cortisol responses, we would expect a strong positive correlation within each day, but we find no significant relationship between self-reported perceived stress and salivary cortisol levels (at baseline, there is a positive association; on the instructional day, the effect is slightly positive; and on exam day, there is a negative association). Given the many confounds influencing salivary cortisol levels, a random effects regression model to control for repeated measures is a more appropriate analysis. This analysis reveals a non-significant association between salivary cortisol and self-reported perceived stress ($\beta = -0.01$; $p > 0.1$). This result indicates that even within individuals, there is no significant association between self-reported perceived stress and salivary cortisol.
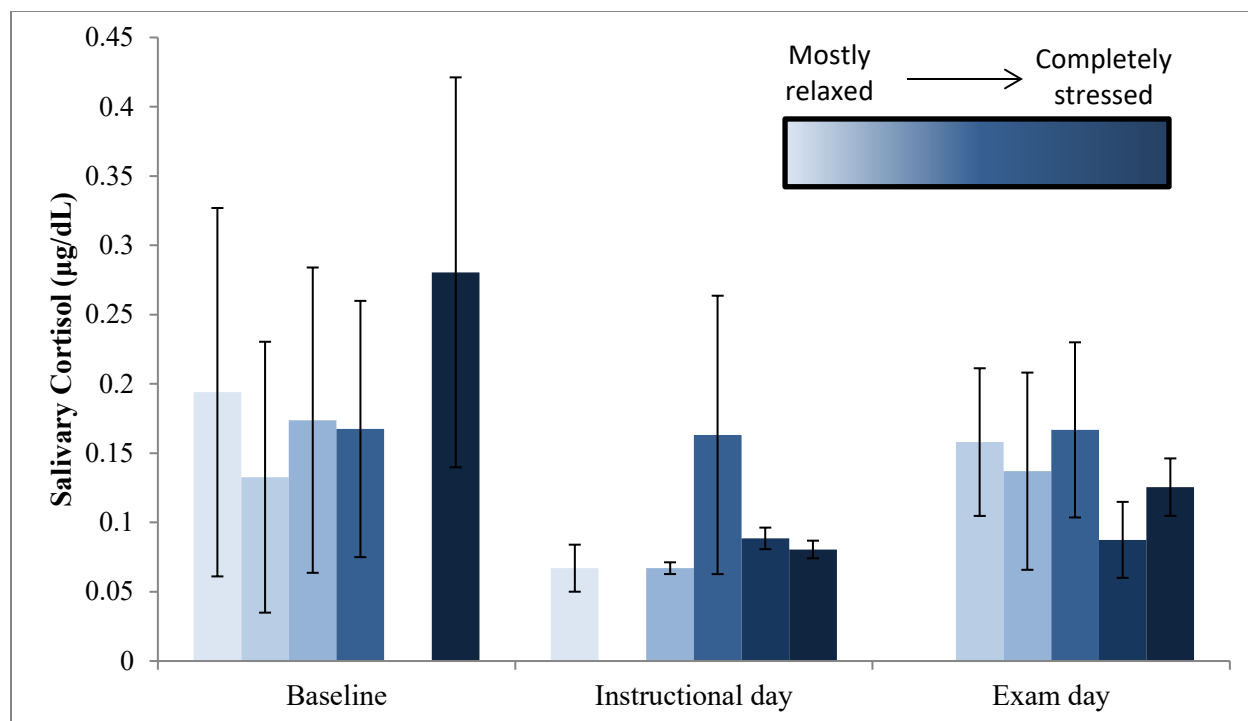
Figure 2: Mean salivary cortisol (µg/dL) by self-reported perceived stress (Lightest blue = mostly relaxed, Darkest blue = completely stress) and day of sample collection (baseline, instructional day, or exam day) with error bars representing +/- one standard deviation. No subject ever reported being *completely relaxed*.

Salivary Testosterone

Figure 3 displays the boxplots of salivary testosterone levels after each of the five conditions. Salivary testosterone after the small-group activity had the lowest mean and standard deviation for both males (M = 107.13 pg/mL; SD = 23.25) and females (M = 55.80 pg/mL; SD = 26.50). Average salivary testosterone was highest after traditional lecture for males (M = 121.93 pg/mL; SD = 29.22) and multiple choice exams for females (M = 108.10 pg/mL; SD = 59.20). The remaining three conditions had similar means. For females, this included: baseline (M = 80.12 pg/mL; SD = 45.52), traditional lecture (M = 70.72 pg/mL; SD = 40.77), and essay exam (M = 82.95 pg/mL; SD = 47.18). For males, this included: baseline (M = 115.39 pg/mL; SD = 46.93), multiple choice exam (M = 115.32 pg/mL; SD = 82.98), and essay exam (M = 110.88 pg/mL; SD = 64.19).
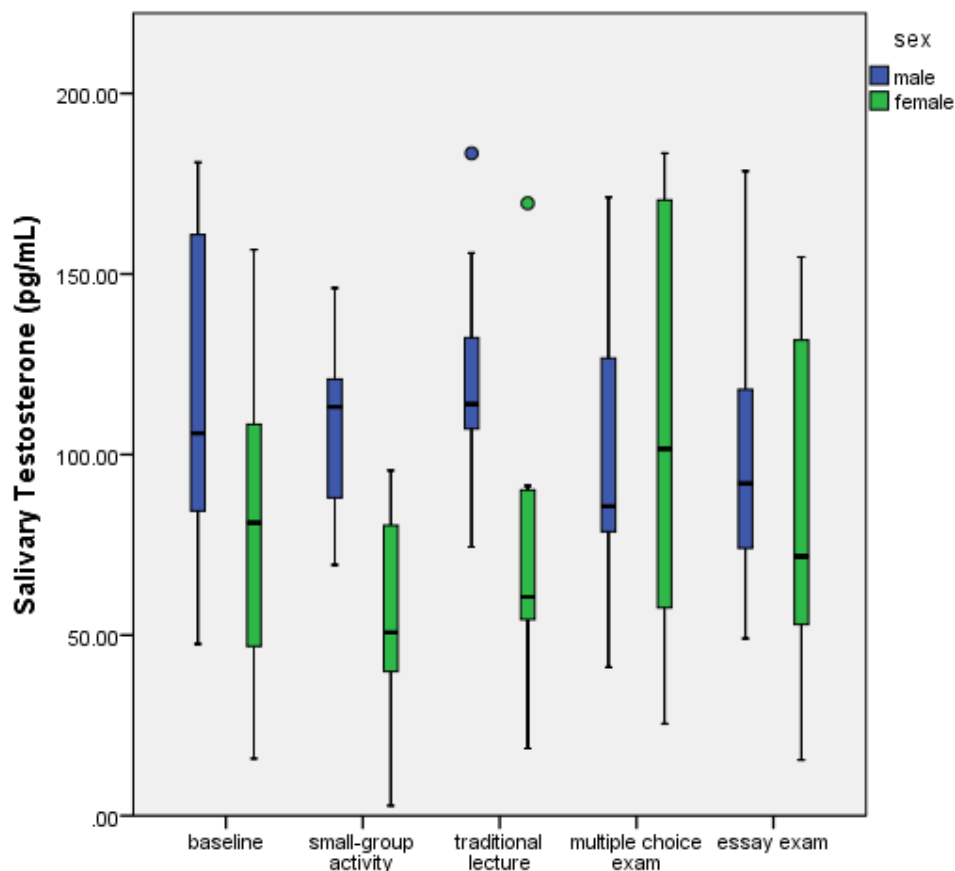
Figure 3: Boxplots of salivary testosterone (measured by pg/mL) across conditions separated by sex of subjects. Boxplots represent the upper quartile, median, and lower quartile as the top, middle, and bottom of the box. The whiskers represent the top 25% (upper whisker) and the bottom 25% (lower whisker) of data, unless one of the scores is greater than the upper quartile plus 1.5 times the inter-quartile range (but less than three times the inter-quartile range), which are deemed outliers and represented by an open circle (o). One male subject had extreme outliers on exam day – after both multiple choice and essay (over 250 pg/mL) and is not displayed in this figure.

We conducted one-way repeated measures ANOVA to compare the effects of small group activity, traditional lecture, multiple choice exam, and essay exam on student's testosterone levels. Mauchly's test indicated that the assumption of sphericity was violated ($p < 0.01$); therefore the Greenhouse-Geisser corrected tests are reported (Field, 2013). These results showed that testosterone levels are *not* significantly different across the five conditions ($p > 0.1$). Table 2 displays the pair-wise comparisons with Šidák correction and Tukey LSD post-hoc tests. Examining our pair-wise comparisons (to determine if groups were significantly different from each other), we found only one significant difference; salivary testosterone was significantly higher following traditional lecture than after the small group activity (Šidák correction, $p < 0.10$; Tukey LSD, $p < 0.01$). There were no significant differences across any other groups.

Table 2: Pairwise comparisons for one-way repeated measures ANOVA for the effect of condition on salivary testosterone

Sample 1	Sample 2	Mean difference	S.E.	*p*-value Šidák correction	*p*-value No correction
Baseline	Small group activity	12.646	8.407	.805	.151
	Lecture	-3.218	9.211	1.000	.731
	Multiple Choice	-12.553	14.128	.992	.387
	Essay	-2.343	9.075	1.000	.799
Small group activity	Lecture	-15.863	5.155	.066	.007**
	Multiple Choice	-25.199	16.331	.782	.141
	Essay	-14.988	12.147	.930	.234
Lecture	Multiple Choice	-9.335	14.739	1.000	.535
	Essay	.875	11.538	1.000	.940
Multiple Choice	Essay	10.210	6.018	.681	.108

Note: * p < 0.05, ** p < 0.01

Discussion

Our results demonstrate that salivary cortisol and testosterone responses vary between students and across different learning and testing environments. Students participating in small group activities had salivary cortisol levels that were: a) significantly lower than at baseline and b) lower than after traditional lecture (but the effect was only significant when *not* adjusting for multiple comparisons). These results suggest that small group activities reduce students' physiological stress compared to baseline and lecture conditions and may be the mechanism by which small group learning activities improve student engagement and academic performance. It was unexpected that baseline salivary cortisol measures were, on average, as high or higher than other conditions (see Limitations section for possible explanations). Salivary testosterone, on the other hand, was lower after the small group learning activity than after traditional lecture. This may suggest that small group activities lead to cooperativeness between group members, as individuals tend to exhibit higher testosterone when having to compete with others, although other interpretations are possible (see Limitations section).

Not only did the group-learning context have the lowest average cortisol levels, but it also showed the least amount of variation between individuals. Our interpretation of this result is that undergraduates are less physiologically stressed by group-learning contexts. In animal studies, predictability, social support, and control over one's environment all contribute to mitigate stress (Sapolsky, 2004). We hypothesize that group learning might be making use of these tactics based on our qualitative participation in and knowledge of the group-learning environment. Students may feel more in control of the pace of learning and the predictability of small group activities. An alternative interpretation may be that more anxious students are able to reduce participation, allowing other group members to take control over the direction of the learning process. Group learning draws on providing social support to reduce stress in the learning experience. Previous research has shown that nursing students who accessed social support perceived it as beneficial in coping with the stress of their academic program (Reeve, Shumaker, Yearwood, Crowell, & Riley, 2013). Future research should explore how small group activities affect students' feeling of control, predictability, and social support, and whether these influence stress and learning.

In our study, participants remained anonymous to the instructor, therefore the test results of each participant cannot be determined and outcome-based measures were not explored, but evidence from prior research has shown that students with higher salivary cortisol before an exam tend to also have significantly lower examination scores (Ng, Koh, & Chia, 2003). Other research has shown that an individual's perceived ability in a subject area is negatively associated with cortisol response to an examination (Minkley, Westerholt, & Kirchner, 2014). While we were unable to explore the association with actual exam scores of our participants, we did collect information on student's anticipated exam score and level of exam preparation (both collected at the end of the exam), and general test-taking anxiety (collected at baseline). Results show that students who reported higher test taking anxiety believed they would earn fewer points on the exam. Students who reported they were more prepared for the exam reported a higher expected exam score. But, there was no significant correlation between these measures (perceived exam performance, exam preparation, or test-taking anxiety) and salivary cortisol after the multiple choice or essay portions of the exam (examined as either cortisol level or change from baseline). This links to our other results examining self-reported stress and cortisol levels, where we found no correlation across any of the learning and exam contexts for self-reported stress and salivary cortisol. This replicates some previous research that has found no correlation between salivary cortisol and self-reported stress (e.g. Murphy et al., 2010; Weekes et al., 2006), but other explanations are possible (see Limitations section). We encourage researchers to keep this in mind when using self-reported measures of stress as their only measure of this socio-biological phenomenon, as it may not be a good proxy of physiological stress.

Variation in hormonal responses across learning and testing contexts most likely responds, in part, to different preferences for learning and evaluation techniques. As always, professors may benefit from using a variety of learning activities in the classroom to reach multiple learning styles (Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010). One of the mechanisms by which small group learning may increase student engagement and improve student outcomes is by reducing stress, increasing control over learning, and allowing classmates to provide each other with social support. Physiological measures of stress vary according to individual physiology and learning type and may play a role in educational outcomes in undergraduate classrooms.

Limitations

While this study is a promising first step to demonstrating the potential positive physiological benefits of small group activities, there are several limitations to our study. First, our subjects were undergraduate students at Boise State University, a largely homogenous group across age, ethnicity, and cultural background. Future research could benefit from investigating patterns of physiological stress in a broader range of learners and learning experiences.

Second, our baseline measures were higher than other conditions and exhibited high variation across subjects. In reflecting on why this may have occurred, it is possible that spitting into a tube in front of peers or an instructor for the first time might have caused participants to feel stressed (particularly among shy or socially anxious subjects (Hofmann, Moscovitch, & Kim, 2006)). Collecting baseline samples within 15 minutes of the beginning of class may also represent the physically and psychologically stressful experience of getting to class (e.g., commuting, pressure to arrive on time (Stutzer & Frey, 2008)) or students consuming food or beverage within an hour of saliva collection (as this information was not collected at baseline). We recommend multiple baseline collections so that students become comfortable with the collection procedure. Providing students with a private location to provide their saliva sample may reduce the stress of spitting in front of peers and faculty.

Third, we found no correlation between self-reported stress level and salivary cortisol. While this replicates some previous research (Hellhammer, Wüst, & Kudielka, 2009; Murphy et al., 2010;

Weekes et al., 2006), it is also possible that our measure of salivary cortisol is measuring some other aspect of physiological stress (e.g., increased cortisol after exercising) or that our self-reported measure of perceived stress (using an ordinal scale) was an imperfect way to capture a person's feeling of stress.

While we did not hypothesize a particular relationship between salivary testosterone and learning or exam contexts, we found that salivary testosterone after small group learning was lower than following lecture. Our research design examined within individual differences in testosterone, but previous research has shown that many factors (some of which vary within individuals) are associated with testosterone levels, including stress, competition, relationship status, parenting status, time of day, exposure to attractive potential partners, gender composition of groups, among others (Booth et al., 1989; Gettler, McDade, Agustin, Feranil, & Kuzawa, 2015; Gray, Kahlenberg, Barrett, Lipson, & Ellison, 2002; Kivlighan et al., 2005; Ronay & Hippel, 2010). Similarly, salivary cortisol can be influenced by many factors, including season, age, exercise, alcohol consumption, education, academic performance measures, biological sex, hormonal birth control, prior trauma, smoking, reproductive state (pre or post-menopausal), medications, chronotype, among others (Badrick, Kirschbaum, & Kumari, 2007; Dowd et al., 2011; Follenius, Brandenberger, Hietter, Simeoni, & Reinhardt, 1982; Hansen et al., 2008; Hellhammer et al., 2009; Lighthall et al., 2013; Persson et al., 2008; Preuß et al., 2010; Stephens et al., 2016; Suzuki et al., 2014; Vgontzas et al., 2003). While our study design compared samples from the same individuals over time to control for within-subject variation, increasing sample size can provide more confidence in the results and may allow for added controls for those factors that influence an individual across time. Additionally, while we tried to control for some of these factors (by excluding participants taking anxiety medication and collecting information on timing of last meal and time since waking), it is possible that students did not accurately report this information or engaged in consumption of beverages/food that they did not consider a meal. Either scenario may have inadvertently influenced their salivary hormones levels, thereby confounding our results.

Conclusion

In conclusion, this study suggests that small group learning activities may reduce salivary cortisol levels, which are linked to a reduction in physiological stress. Group learning activities may lead to improved learning outcomes by mitigating students' physiological distress. Although we found no evidence connecting self-reported stress to our physiological measures, small groups may be effective at mitigating stress by increasing control and predictability of the learning environment while adding social support to the learning process, allowing small group activities to result in improved student engagement and academic performance. This research suggests that small group learning may not just improve academic performance; it may also contribute to reduced physiological stress and associated positive health benefits.

Acknowledgements

References

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress*, *6*, 113–126. http://doi.org/10.1016/j.ynstr.2016.11.001

Ambrose, S., Bridges, M., DiPietro, M., Lovett, M., & Norman, M. (2010). *How learning works: Seven research-based principles for smart teaching*. San Francisco: Jossey-Bass.

Badrick, E., Kirschbaum, C., & Kumari, M. (2007). The relationship between smoking status and cortisol secretion. *Journal of Clinical Endocrinology and Metabolism*, *92*(3), 819–824. http://doi.org/10.1210/jc.2006-2155

Bardi, M., Koone, T., Mewaldt, S., & O'Connor, K. (2011). Behavioral and physiological correlates of stress related to examination performance in college chemistry students. *Stress*, *14*(5), 557–566. http://doi.org/10.3109/10253890.2011.571322

Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone, and winning and losing in human competition. *Hormones and Behavior*, *23*(4), 556–571. https://doi.org/10.1016/0018-506X(89)90042-1

Bradford, J., Mowder, D., & Bohte, J. (2016). You can lead students to water, but you can't make them think: An assessment of student engagement and learning through Student-Centered Teaching. *Journal of the Scholarship of Teaching and Learning*, *16*(4), 33–43. http://doi.org/10.14434/josotl.v16i4.20106

Byun, C. H. C. (2014). The prisoner's dilemma and economics 101: Do active learning exercises correlate with student performance? *Journal of the Scholarship of Teaching and Learning*, *14*(5), 79. http://doi.org/10.14434/josotlv14i5.12950

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: Does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology*, *37*(8), 1111–1134. http://doi.org/10.1016/j.psyneuen.2011.12.010

Coakley, L. A., & Sousa, K. J. (2013). The effect of contemporary learning approaches on student perceptions in an introductory business course. *Journal of the Scholarship of Teaching and Learning*, *13*(3), 1–22.

Dowd, J. B., Ranjit, N., Do, D. P., Young, E. A., House, J. S., & Kaplan, G. A. (2011). Education and levels of salivary cortisol over the day in US adults. *Annals of Behavioral Medicine*, *41*(1), 13–20. http://doi.org/10.1007/s12160-010-9224-2

Ellison, P. T., & Gray, P. B. (2009). *Endocrinology of social relationships*. Cambridge, MA & London, England: Harvard University Press.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Los Angeles: Sage Publications.

Flegr, J., & Priplatova, L. (2010). Testosterone and cortisol levels in university students reflect actual rather than estimated number of wrong answers on written exam. *Neuroendocrinology Letters*, *31*(4), 577–581.

Follenius, M., Brandenberger, G., Hietter, B., Simeoni, M., & Reinhardt, B. (1982). Diurnal cortisol peaks and their relationship to meals. *Journal of Clinical Endocrinology and Metabolism*, *55*(4), 757–761. https://doi.org/10.1210/jcem-55-4-757

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–5. http://doi.org/10.1073/pnas.1319030111

Gettler, L. T., McDade, T. W., Agustin, S. S., Feranil, A. B., & Kuzawa, C. W. (2015). Longitudinal perspectives on fathers' residence status, time allocation, and testosterone in the Philippines. *Adaptive Human Behavior and Physiology*, *1*(2), 124–149. http://doi.org/10.1007/s40750-014-

0018-9

Gray, P. B., Kahlenberg, S. M., Barrett, E. S., Lipson, S. F., & Ellison, P. T. (2002). Marriage and fatherhood are associated with lower testosterone in males. *Evolution and Human Behavior*, *23*(3), 193–201. http://doi.org/10.1016/S1090-5138(01)00101-5

Hansen, A. M., Garde, A. H., & Persson, R. (2008). Sources of biological and methodological variation in salivary cortisol and their impact on measurement among healthy adults: a review. *Scandinavian Journal of Clinical and Laboratory Investigation*, *68*(6), 448–58. http://doi.org/10.1080/00365510701819127

Harden, K. P., Wrzus, C., Luong, G., Grotzinger, A., Bajbouj, M., Rauers, A., … Riediger, M. (2016). Diurnal coupling between testosterone and cortisol from adolescence to older adulthood. *Psychoneuroendocrinology*, *73*, 79–90. http://doi.org/10.1016/j.psyneuen.2016.07.216

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, *34*(2), 163–171. http://doi.org/10.1016/j.psyneuen.2008.10.026

Hofmann, S. G., Moscovitch, D. A., & Kim, H. J. (2006). Autonomic correlates of social anxiety and embarrassment in shy and non-shy individuals. *International Journal of Psychophysiology*, *61*(2), 134–142. http://doi.org/10.1016/j.ijpsycho.2005.09.003

Hoyt, L. T., Zeiders, K. H., Ehrlich, K. B., & Adam, E. K. (2016). Positive upshots of cortisol in everyday life. *Emotion*, *16*(4), 431–435. http://doi.org/10.1037/emo0000174

Kenwright, K., Liddell, P. W., Bloom, L., Zucker-Levin, A., Nolen, A. H., Faulkner, L. W., & Batorski, R. E. (2011). Salivary cortisol levels in students challenged with a testing stressor. *Clinical Laboratory Science*, *24*(4), 221–6.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'- A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*, 76–81. http://doi.org/119004

Kivlighan, K. T., Granger, D. A., & Booth, A. (2005). Gender differences in testosterone and cortisol response to competition. *Psychoneuroendocrinology*, *30*(1), 58–71. http://doi.org/10.1016/j.psyneuen.2004.05.009

Lacey, K., Zaharia, M. D., Griffiths, J., Ravindran, A. V., Merali, Z., & Anisman, H. (2000). A prospective study of neuroendocrine and immune alterations associated with the stress of an oral academic examination among graduate students. *Psychoneuroendocrinology*, *25*, 339–356. http://doi.org/10.1016/S0306-4530(99)00059-1

Lee, B. K., Glass, T. A., McAtee, M. J., Wand, G. S., Bandeen-Roche, K., Bolla, K. I., & Schwartz, B. S. (2007). Associations of salivary cortisol with cognitive function in the Baltimore memory study. *Archives of General Psychiatry*, *64*(7), 810–8. http://doi.org/10.1001/archpsyc.64.7.810

Leproult, R., Copinschi, G., Buxton, O., & Van Cauter, E. (1997). Sleep loss results in an elevation of cortisol levels the next evening. *Sleep: Journal of Sleep Research & Sleep Medicine*, *20*(10), 865–870. https://doi.org/10.1093/sleep/20.10.865

Lighthall, N. R., Gorlick, M. A., Schoeke, A., Frank, M. J., & Mather, M. (2013). Stress modulates reinforcement learning in younger and older adults. *Psychology and Aging*, *28*(1), 35–46. http://doi.org/10.1037/a0029823

Martinek, L., Oberascher-Holzinger, K., Weishuhn, S., Klimesch, W., & Kerschbaum, H. H. (2003). Anticipated academic examinations induce distinct cortisol responses in adolescent pupils. *Neuroendocrinology Letters*, *24*(6), 449–453.

Mazur, A., Booth, A., & Dabbs, J. M. (1992). Testosterone and chess competition. *Social Psychology Quarterly*, *55*(1), 70–77. http://doi.org/10.2307/2786687

Mazur, A., Susman, E. J., & Edelbrock, S. (1997). Sex difference in testosterone response to competition in a video game. *Evolution and Human Behavior*, *18*, 317–326.

https://doi.org/10.1016/S1090-5138(97)00013-5

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *The New England Journal of Medicine*, *338*(3), 171–179. http://doi.org/10.1056/NEJM199801153380307

Merrifield, C., & Danckert, J. (2014). Characterizing the psychophysiological signature of boredom. *Experimental Brain Research*, *232*(2), 481–491. http://doi.org/10.1007/s00221-013-3755-2

Merz, C. J., & Wolf, O. T. (2015). Examination of cortisol and state anxiety at an academic setting with and without oral presentation. *Stress*, *18*(1), 138–142. https://doi.org/10.3109/10253890.2014.989206

Minkley, N., Westerholt, D. M., & Kirchner, W. H. (2014). Academic self-concept of ability and cortisol reactivity. *Anxiety, Stress, and Coping*, *27*(3), 303–16. http://doi.org/10.1080/10615806.2013.848273

Moffat, K. J., McConnachie, A., Ross, S., & Morrison, J. M. (2004). First year medical student stress and coping in a problem-based learning medical curriculum. *Medical Education*, *38*(5), 482–491. http://doi.org/10.1046/j.1365-2929.2004.01814.x

Moog, R. S., Creegan, F. J., Hanson, D. M., Spencer, J. N., & Straumanis, A. R. (2006). Process-Oriented Guided Inquiry Learning: POGIL and the POGIL project. *Metropolitan Universities Journal*, *17*(4), 41–52.

Murphy, L., Denis, R., Ward, C. P., & Tartar, J. L. (2010). Academic stress differentially influences perceived stress, salivary cortisol, and immunoglobulin-A in undergraduate students. *Stress*, *13*(4), 365–370. http://doi.org/10.3109/10253891003615473

Ng, V., Koh, D., & Chia, S. E. (2003). Examination stress, salivary cortisol, and academic performance. *Psychological Reports*, *93*(3), 1133–1134. http://doi.org/http://dx.doi.org/10.2466/PR0.93.8.1133-1134

Persson, R., Garde, A. H., Hansen, A. M., Osterberg, K., Larsson, B., Orbaek, P., & Karlson, B. (2008). Seasonal variation in human salivary cortisol concentration. *Chronobiology International*, *25*(6), 923–937. http://doi.org/10.1080/07420520802553648

Preuß, D., Schoofs, D., Schlotz, W., & Wolf, O. T. (2010). The stressed student: Influence of written examinations and oral presentations on salivary cortisol concentrations in university students. *Stress*, *13*(3), 221–229. http://doi.org/10.3109/10253890903277579

Reeve, K. L., Shumaker, C. J., Yearwood, E. L., Crowell, N. A., & Riley, J. B. (2013). Perceived stress and social support in undergraduate nursing students' educational experiences. *Nurse Education Today*, *33*(4), 419–424. http://doi.org/10.1016/j.nedt.2012.11.009

Robinson, S. J., Sünram-Lea, S. I., Leach, J., & Owen-Lynch, P. J. (2008). The effects of exposure to an acute naturalistic stressor on working memory, state anxiety and salivary cortisol concentrations. *Stress*, *11*(2), 115–124. http://doi.org/10.1080/10253890701559970

Ronay, R., & Hippel, W. V. (2010). The presence of an attractive woman elevates testosterone and physical risk taking in young men. *Social Psychological and Personality Science*, *1*(1), 57–64. http://doi.org/10.1177/1948550609352807

Salimetrics. (2014). Expanded Range High Sensitivity Salivary Cortisol Enzyme Immunoassay Kit. Retrieved October 17, 2017, from https://www.salimetrics.com/assets/documents/1-3002n.pdf

Salimetrics. (2016). Expanded Range Salivary Testosterone Enzyme Immunoassay Kit. Retrieved July 10, 2016, from https://www.salimetrics.com/assets/documents/1-2312n.pdf

Sapolsky, R. M. (2004). *Why zebras don't get ulcers* (3rd ed.). New York, NY: Henry, Holt & Company, Inc.

Schoofs, D., Hartmann, R., & Wolf, O. T. (2008). Neuroendocrine stress responses to an oral academic examination: No strong influence of sex, repeated participation and personality traits. *Stress*, *11*(1), 52–61. https://doi.org/10.1080/10253890701453943

Simonson, S. R., & Shadle, S. E. (2013). Implementing Process Oriented Guided Inquiry Learning (POGIL) in undergraduate biomechanics: Lessons learned by a novice. *Journal of STEM Education*, *14*(1), 56–64.

Singh, R., Goyal, M., Tiwari, S., Ghildiyal, A., Nattu, S. M., & Das, S. (2012). Effect of examination stress on mood, performance and cortisol levels in medical students. *Indian Journal of Physiology and Pharmacology*, *56*(1), 48–55.

Stephens, M. A. C., Mahon, P. B., McCaul, M. E., & Wand, G. S. (2016). Hypothalamic-pituitary-adrenal axis response to acute psychosocial stress: Effects of biological sex and circulating sex hormones. *Psychoneuroendocrinology*, *66*, 47–55. http://doi.org/10.1016/j.psyneuen.2015.12.021

Stutzer, A., & Frey, B. S. (2008). Stress that doesn't pay: The commuting paradox. *Scandinavian Journal of Economics*, *110*(2), 339–366. http://doi.org/10.1111/j.1467-9442.2008.00542.x

Suzuki, A., Poon, L., Papadopoulos, A. S., Kumari, V., & Cleare, A. J. (2014). Long term effects of childhood trauma on cortisol stress reactivity in adulthood and relationship to the occurrence of depression. *Psychoneuroendocrinology*, *50*, 289–299. http://doi.org/10.1016/j.psyneuen.2014.09.007

Swap, R. J., & Walter, J. A. (2015). An approach to engaging students in a large-enrollment, introductory STEM college course. *Journal of the Scholarship of Teaching and Learning*, *15*(5), 1–21. http://doi.org/10.14434/josotl.v16i5.18910

Takatsuji, K., Sugimoto, Y., Ishizaki, S., Ozaki, Y., Matsuyama, E., & Yamaguchi, Y. (2008). The effects of examination stress on salivary cortisol, immunoglobulin A, and chromogranin A in nursing students. *Biomedical Research*, *29*(4), 221–224. http://doi.org/10.2220/biomedres.29.221

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*(3), 411–429. http://doi.org/10.1037/0033-295X.107.3.411

Vedhara, K., Hyde, J., Gilchrist, I. D., Tytherleigh, M., & Plummer, S. (2000). Acute stress, memory, attention and cortisol. *Psychoneuroendocrinology*, *25*(6), 535–549. http://doi.org/10.1016/S0306-4530(00)00008-1

Vgontzas, A. N., Zoumakis, M., Bixler, E. O., Lin, H. M., Prolo, P., Vela-Bueno, A., … Chrousos, G. P. (2003). Impaired nighttime sleep in healthy old versus young adults is associated with elevated plasma interleukin-6 and cortisol levels: Physiologic and therapeutic implications. *Journal of Clinical Endocrinology and Metabolism*, *88*(5), 2087–2095. http://doi.org/10.1210/jc.2002-021176

Villanueva, I., Valladares, M., & Goodridge, W. (2016). Use of galvanic skin responses, salivary biomarkers, and self-reports to assess undergraduate student performance during a laboratory exam activity. *Journal of Visualized Experiments: JoVE*, (108), e53255. http://doi.org/10.3791/53255

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE - Life Sciences Education*, *7*(1), 361–367. http://doi.org/10.1187/cbe.08

Walker, L., & Warfa, A. R. M. (2017). Process oriented guided inquiry learning (POGIL®) marginally effects student achievement measures but substantially increases the odds of passing a course. *PLoS ONE*, *12*(10), 1–17. http://doi.org/10.1371/journal.pone.0186203

Weekes, N., Lewis, R., Patel, F., Garrison-Jakel, J., Berger, D. E., & Lupien, S. J. (2006). Examination stress as an ecological inducer of cortisol and psychological responses to stress in undergraduate students. *Stress*, *9*(4), 199–206. http://doi.org/10.1080/10253890601029751

The Funny Thing Is, Instructor Humor Style Affects Likelihood of Student Engagement

Kristie Nienaber
Nienaber Counseling

Gwyneth Abrams
Canadian Coast Guard

Dan Segrist
Southern Illinois University Edwardsville
dsegris@siue.edu

Abstract: Instructors often use humor in teaching their classes. Research suggests that humor can affect how instructors and their teaching are perceived. The current study examined whether the type of humor used by a hypothetical instructor and instructor gender affected the perceived likelihood of engaging with the instructor. College students read a vignette describing the teaching and humor used by a hypothetical instructor. The likelihood that students would engage with the instructor was highest when the instructor's humor style was good-natured and lowest when it was hostile and sarcastic. Instructor gender had no effect on students' likelihood of engaging with the instructor.

Keywords: humor, student engagement

Instructors are often encouraged to incorporate humor into their teaching and frequently do so (Martin, 2007; Torok, McMorris, & Linn, 2004). Multiple authors have reviewed the literature on the use and effect of humor in the college classroom (see Banas, Dunbar, Rodriguez, & Liu, 2011; Martin, 2007; Martin, Preiss, Gayle, & Allen, 2006; Oppliger, 2003; Powers, 2008; Segrist & Hupp, 2015). Proponents of the use of humor suggest that humor can have positive effects on students, the retention of material, and the learning environment (e.g., Berk, 2002; Deitar, 2000; Garner, 2006). Torok, McMorris, and Lin (2004) point out that, "Humor, appropriately used, has the potential to humanize, illustrate, defuse, encourage, reduce anxiety, and keep people thinking" (p. 14).

Professors who incorporate humor in their teaching may be more positively evaluated by students (e.g., Bryant, Comisky, Crane, & Zillmann, 1980; Garner, 2006; Van Giffen, 1990; Wanzer & Frymier, 1999). Increased use of humor has also been associated with increased instructor immediacy (Gorham & Christophel, 1990; Wanzer & Frymier, 1999). Wilson and Taylor (2001) found that instructor humor positively related to students' perceptions that instructor "wanted them to succeed" and had a "genuine concern" for and "positive attitude toward" them (pp. 136-137). Instructor humor can even affect students' willingness to engage an instructor outside of class. Aylor and Oppliger (2003) found that instructors' use of humor in the classroom predicted the frequency of students' out-of-class communication with those instructors. More recently, Goodboy and colleagues (2015) demonstrated that college students' ratings of the frequency and effectiveness of an instructor's humor predicted both in-class participation and out-of-class communication with that instructor.

Hackathorn and colleagues (2011) found that when instructors used humor to teach material, students demonstrated better knowledge and comprehension of the material than when instructors did not use humor. Although research evidence demonstrates that humor can augment teaching efforts, it is not a magic bullet. In fact, the effects of instructor humor are not always positive or appreciated. For example, in general students consider it inappropriate for teachers to use hostile

humor, disparaging humor, sexual humor, or humor that targets students individually or collectively (e.g., humor targeting sex, sexual orientation, ethnicity, or religion; Torok, McMorris, & Lin, 2004; Wanzer, Frymier, Wojtaszczyk, & Smith, 2006). Furthermore, hostile humor can create an adverse classroom milieu (Stuart & Rosenfeld, 1994).

Research on the use of humor in teaching has suggested that students' perceptions of instructor humor, as well as the effect of that humor, may be influenced by the gender of the instructor. For example, Van Giffen (1990) found that for female instructors, but not male instructors, the use of humor was predictive of perceived instructor effectiveness. Gorham and Christophel (1990), however, found that although male and female instructors in their study used similar forms of humor, only use of humor by male instructors related to students' self-reported learning. Bryant, Commiskey, Crane, and Zillman (1980) found that the use of humor was positively related to the perceived effectiveness and appeal of male instructors. For female instructors, no forms of humor were positively related to perceived effectiveness but some (i.e., puns, distracted humor) were negatively associated with perceived effectiveness and appeal. Hostile humor was positively related to the perceived effectiveness of male instructors while aggressive and sexual/hostile humor were associated with increased appeal for female instructors. In her review of the literature on humor and learning, Oppliger (2003) called for more research examining humor in instructor gender and the use of humor in teaching.

The present study examined the effect of the type of humor used by a hypothetical professor on students' perceived likelihood of engaging with that professor. Because previous research suggests that the perception of instructor humor may be affected by instructor gender as well as the interaction of instructor gender and humor type, the present study included both humor type and instructor gender as variables. We also speculated that introversion might affect students' comfort with interacting with a professor in or out of class, (regardless of that professor's use of humor). As a result, we included introversion as a control variable.

Method

Participants

The sample included 157 undergraduate students attending a Midwestern public university and was comprised primarily of freshmen (53.5%) and sophomores (27.4%) with a mean age of 19.79 years ($SD = 2.58$). The majority of participants identified as female (75.2%) and ethnically as Caucasian (61.1%) or African American (32.5%).

Materials and Procedure

Participants who signed up for this study were provided with a Survey Monkey link through which they could access the study. After providing consent, participants completed demographic questions. In order to randomly assign vignette condition, participants also indicated their birth month. Based on birth month, the survey directed participants to one of six possible vignettes briefly describing a hypothetical college professor and his or her teaching style and use of humor in the classroom. Subsequently participants rated a series of items about the likelihood of engaging and comfort level with the instructor. Participants also completed an introversion measure.

Vignettes

The vignettes (Appendix 1) briefly described "Professor Morgan" who is teaching *Psychology 101*. The vignette included information about Professor Morgan's teaching style and use of humor in teaching. Six vignettes were created and varied by professor gender and the professor's use of humor in teaching (i.e., affiliative humor, aggressive humor, little/no humor). For the vignettes that depicted Professor Morgan using humor in teaching, descriptions of the professor's style of humor were drawn from items comprising two of the four subscales included in the Humor Styles Questionnaire (Martin et al., 2003). Affiliative humor is, "an essentially non-hostile, tolerant form of humor that is affirming of others and presumably enhances interpersonal attraction and cohesion" (Martin et al., 2003, p. 53). Aggressive humor includes sarcasm and using humor to "ridicule and manipulate others" (Martin, 2003, p. 23). As can be seen in Appendix A, the "little/no" humor vignette is shorter than the "affiliative" and "aggressive" humor vignettes. Although all three vignettes would ideally be the same length, we felt it was important to keep the content of the vignettes focused on the humor style.

Engaging with Professor Scale

Participants' likelihood of engaging and comfort level with the hypothetical professor was assessed through nine items based loosely on original item pool for the Professor-Student Rapport Scale (Wilson, Ryan, & Pugh, 2010) and tailored to the vignette character (i.e., Professor Morgan; see Table 1). Items encompassed potential student behaviors and attitudes reflecting willingness to engage and comfort level with the instructor and the instructor's class (e.g., *I would feel comfortable asking a question during class; I would enjoy having this professor; I would participate in class discussions in Professor Morgan's class*) as well as the likelihood of contacting the professor outside of class (e.g., *I would go to Professor Morgan's office hours; I would email Professor Morgan with a question or concern about class; I would feel comfortable letting Professor Morgan know I need help with class material*). Items were rated on a five-point Likert scale (1=strongly disagree…5=strongly agree). Total scores could range from 9 to 45, with higher scores reflecting increased likelihood of engagement. Cronbach's alpha for the scale was .94.

Table 1: Items comprising the Engaging with Professor Scale

I would go to Professor Morgan's office hours
I would participate in class discussions in Professor Morgan's class
I would be motivated to attend Professor Morgan's class
I would feel comfortable in Professor Morgan's class
I would email Professor Morgan with a question or concern about class
I would feel comfortable letting Professor Morgan know I need help with class material
I would point out a mistake on Professor Morgan's slides/lecture during class
I would feel comfortable asking a question during class
I would enjoy having this professor

Introversion

Introversion was assessed through a 10-item International Personality Item Pool (Goldberg, et al., 2006) scale based on the California Psychological Inventory (Gough & Bradley, 1996). In the present study, Cronbach's alpha for the scale was .75.

Results

We intended to analyze data using a 2 X 3 between groups ANCOVA in order to control for participant introversion; however, there was not a significant relationship between introversion scores and the perceived comfort level with the professor ($r = .04$, $p = .604$). Consequently, introversion was not used as a covariate and a 2 X 3 between groups ANOVA was conducted to assess the effect of professor humor style and professor gender on students' comfort level and likelihood of engaging with the professor. The independent variables were gender of the hypothetical professor and style humor of used - affiliative, aggressive, or little/no humor. The dependent variable was likelihood of engaging with the professor.

The interaction between professor gender and humor style was not significant, $F(2, 146) = .65$, $p = .525$. The main effect for instructor humor style was significant, $F(2, 146) = 71.37$, $p < .01$, $\eta^2 = .49$. Post-hoc comparisons using the Tukey HSD test indicated significant differences at the $p < .05$ level among all three humor conditions (Figure 1). Students were more likely to engage with a professor portrayed as using affiliative humor ($M = 38.10$, $SD = 5.61$) compared to little/no humor ($M = 26.19$, $SD = 6.27$) and aggressive humor ($M = 21.70$, $SD = 7.70$). There was not a significant main effect for professor gender, $F(1, 146) = .11$, $p = .738$.
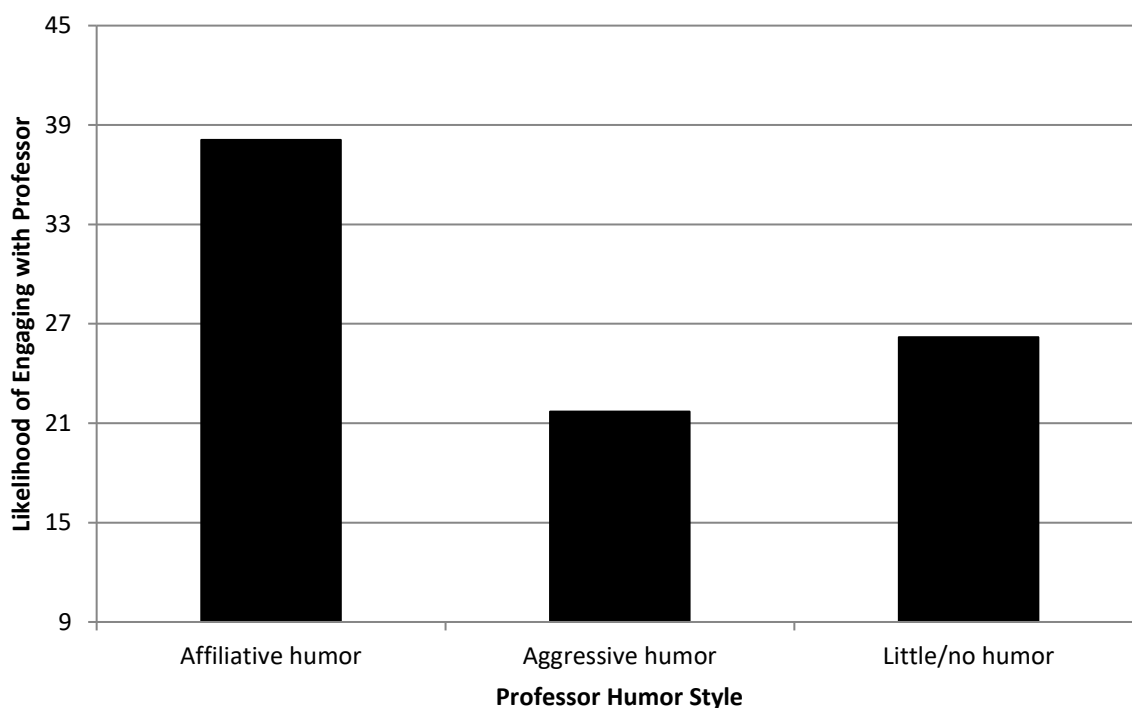
Figure 1: Likelihood of Engaging with Professor based on Type of Humor Used

Discussion

The results of the present study suggest that students are more likely to engage and feel comfortable with an instructor who uses affiliative humor as opposed to hostile humor or no humor at all. The difference between the likelihood of engagement ratings for affiliative humor and aggressive humor or no humor are striking. In the affiliative humor condition, comfort ratings are well above the scale median (almost two standard deviations above the median). Research has demonstrated a relationship

between instructor humor and immediacy (e.g., Gorham & Christophel, 1990; Wanzer & Frymier, 1999; Wilson & Taylor, 2001). The present study extends this research by demonstrating that different types of humor have different effects on the likelihood of engaging with a professor. Specifically, students were more likely to engage with a professor whose classroom humor was affiliative. This type of humor is a non-threatening form of humor that fosters social bonding (Martin, 2003). Yip and Martin (2006) found that affiliative humor was positively associated with social competence, more specifically initiating relationships and disclosing personal information. Although further research would be needed to test this assertion, the present results suggest that students may have associated the use of affiliative humor with an increased level of social competence and, consequently, imagined feeling more comfortable with "Professor Morgan".

Findings also indicated that students felt more comfortable engaging a professor who used no humor in the classroom than one who used aggressive humor. In fact, in the aggressive humor condition, ratings of perceived comfort with "Professor Morgan" were well below the median of the scale. Wanzer, Frymier, and Irwin (2010) found that a professor's use of disparaging humor was not related to student learning. However, research also indicates that students perceive instructor humor that demeans students as inappropriate (Frymier, Wanzer, & Wojtaszczyk, 2008). Although humor that ridicules might not adversely affect learning, the results of the present study suggest that disparaging humor affects students' comfort with a professor and could, consequently, become a barrier preventing students from seeking help from an instructor. Yip and Martin (2006) found that the use of aggressive humor was negatively associated with emotional perception, emotional support, and conflict management. Perhaps students' low likelihood of engaging with an instructor who used aggressive humor reflects students' assumptions that such an instructor would lack empathy or understanding. Students might understandably be less comfortable approaching or interacting with a professor who uses humor to ridicule and does not seem emotionally supportive.

Interestingly, instructor gender did not interact with type of humor to influence students' comfort level with "Professor Morgan". Much of the research on gender and instructor humor is somewhat dated. It may be that perceptions of instructor humor are not affected by an instructor's gender to the degree or in the same way they were 20 to 30 years ago. However, we concur with Oppliger's (2003) suggestion that more research examining the interaction of instructor gender and instructor humor is warranted.

The current study has a few important limitations. In this study participants were asked to rate items asking about their comfort level with a hypothetical professor. The vignettes provide little information beyond professor gender, course, general teaching style, and type of humor used while teaching. This means that participants lacked other contextual cues (e.g., the specific jokes, delivery, situation). Humor, however, is largely a social experience (Martin, 2007) that is context dependent (Lampert & Ervin-Tripp, 1998 as cited in Crawford, 2003). For example, people tend to laugh more often when they are with others (Martin & Kuiper, 1999). Because participants read about a professor's style of humor, but did not experience an actual professor using humor in a classroom setting the present results should be interpreted with some caution. Future studies might address this methodological issue by showing groups of participants video clips of a professor using humor while teaching.

The majority of participants in this study were first year college students. As a result, their exposure to college professors was fairly limited. As students matriculate they may experience a greater variety of instructors as well as a greater diversity in how humor is used in the classroom. Consequently, the influence of instructor humor on students may change as the student becomes a more seasoned college student. Future studies could explore whether the effect of instructor humor varies as a function of students' matriculation level.

Appendix

Appendix 1. Vignettes.

<u>**No Humor**</u>
Professor Morgan teaches *Psychology 101*. Colleagues would describe her as being highly knowledgeable in course subject matter. During lectures, Professor Morgan is very serious. For example, she rarely if ever makes a joke during class. In general, Dr. Morgan can be described as serious and thorough.

<u>**Aggressive Humor**</u>
Professor Morgan teaches *Psychology 101*. Colleagues and students would describe her as being highly knowledgeable in the course subject matter. During lectures, Professor Morgan is very thorough and uses a lot of humor to maintain students' attention. For example, if a student makes a mistake in class, Professor Morgan often teases them about it, without showing concern about how others are taking her jokes. If something is really funny she will not hesitate to laugh and joke about it, even if someone might be offended. At times Professor Morgan also uses humor at the expense of others. Professor Morgan teases and puts down students she dislikes. Some students are offended by her sense of humor. In general, Dr. Morgan's humor would be considered sarcastic, insensitive, and at times offensive.

<u>**Affiliative Humor**</u>
Professor Morgan teaches *Psychology 101*. Colleagues and students would describe her as being highly knowledgeable in the course subject matter. During lectures, Professor Morgan is very thorough and uses a lot of humor to maintain students' attention. For example, Professor Morgan often jokes around and laughs with other students during class. Most students consider her to be a naturally humorous, witty, and an amusing person. During lectures, Dr. Morgan uses humor that revolves around funny stories in her life. Dr. Morgan finds enjoyment in making students, colleagues, and friends laugh. In general, Dr. Morgan's humor would be considered light-hearted, amusing, charming, and cheerful.

References

Aylor, B., & Opplinger, P. (2003). Out-of-class communication and student perceptions of instructor humor orientation and socio-communicative style. *Communication Education*, *52*(2), 122-134. https://doi.org/10.1080/03634520302469

Banas, J. A., Dunbar, N., Rodriguez, D., & Liu, S. J. (2011). A review of humor in educational settings: Four decades of research. *Communication Education*, *60*, 115-144. https://doi.org/10.1080/03634523.2010.496867

Berk, R. A. (2002). *Humor as an instructional defibrillator: Evidence-based techniques in teaching and assessment*. Sterling, VA: Stylus.

Bryant, J., Commiskey, P. W., Crane, J. S., & Zillmann, D. (1980). Relationship between college teachers' use of humor in the classroom and students' evaluations of their teachers. *Journal of Educational Psychology*, *72*, 511-519. https://doi.org/10.1037//0022-0663.72.4.511

Crawford, M. (2003). Gender and humor in social context. *Journal of Pragmatics*, *35*, 1413-1430. https://doi.org/10.1016/s0378-2166(02)00183-2

Dieter, R. (2000). The use of humor as a teaching tool in the college classroom. *NACTA Journal*, *44*, 20-28.

Frymier, A. B., Wanzer, M. B., & Wojtaszczyk, A. M. (2008). Assessing students' perceptions of inappropriate and appropriate teacher humor. *Communication Education*, *57*(2), 266-288.

https://doi.org/10.1080/03634520701687183

Garner, R. L. (2006). Humor in pedagogy: How ha-ha can lead to aha! *College Teaching*, *54*(1), 177-180. https://doi.org/10.3200/ctch.54.1.177-180

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*, 84-96. https://doi.org/10.1016/j.jrp.2005.08.007

Goodboy, K. A., Booth-Butterfield, M., Bolkan, San, & Griffin, D. J. (2015). The role of instructor humor and students' educational orientations in student learning, extra effort, participation, and out-of-class communication. *Communication Quarterly*, *63*(1), 44-61. https://doi.org/10.1080/01463373.2014.965840

Gorham, J., & Christophel, D. M. (1990). The relationship of teachers' use of humor in the classroom to immediacy and student learning. *Communication Education*, *39*, 46-62. https://doi.org/10.1080/03634529009378786

Gough, H. G., & Bradley, P. (1996). *CPI Manual*, 3rd Edition. Palo Alto, CA: Consulting Psychologists Press.

Hackathorn, J., Garczynski, A. M., Blankmeyer, K., Tennial, R. D., & Solomon, E. (2011). All kidding aside: Humor increases learning at knowledge and comprehension levels. *Journal of the Scholarship of Teaching and Learning*, *11*(4), 116-123.

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/). Internet Web Site.

Lampert, M., & Ervin-Tripp, S. (1998). Exploring paradigms: The study of gender and sense of humor near the end of the 20th century. In Willibald Ruch (Ed.), *The sense of humor,* pp. 231-270. New York: Mouton de Gruyter.

Martin, R. (2007). *The psychology of humor: An integrative approach*. London: Elsevier Academic Press.

Martin, R., & Kuiper, N. A. (1999). Daily occurrence of laughter: Relationships with age, gender, and Type A personality. *Humor: The International Journal of Humor Research*, *12*(4), 355-384. https://doi.org/10.1515/humr.1999.12.4.355

Martin, D. M., Preiss, R. W., Gayle, B. M., & Allen, M. (2006). A meta-analytic assessment of the effect of humorous lectures on learning. In B. M. Gayle, R. W. Preiss, N. Burrell, & M. Allen (Eds.), *Classroom communication and instructional processes: Advances through meta-analysis*. (pp. 295-313). Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in the uses of humor and their relation to psychology well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*, *37*, 48-75. https://doi.org/10.1016/s0092-6566(02)00534-2

Oppliger, P. A. (2003). Humor and learning. In J. Bryant, D. Roskos-Ewoldsen, & J. Cantor (Eds.), *Communication and emotion: Essays in honor of Dolf Zillmann* (pp. 255-273). New York: Routledge.

Powers, T. (2008). *Engaging students with humor*. In B. Perlman, L. I. McCann, & S. H. McFadden (Eds.), *Lessons Learned, Volume 3: Practical advice for the teaching of psychology* (pp. 53-62). Washington, DC: Association for Psychological Science.

Segrist, D. J., & Hupp, Stephen, D. A. (2015). *This class is a joke: Humor as a pedagogical tool in the teaching of psychology*. Retrieved from http://teachpsych.org/Resources/Documents/otrp/resources/segrist15.pdf

Stuart, W. D., & Rosenfeld, L. B. (1994). Student perceptions of teacher humor and classroom climate. *Communication Research Reports*, *11*(1), 87-97. https://doi.org/10.1080/08824099409359944

Torok, S. E., McMorris, R. F., & Lin, W. C. (2004). Is humor an appreciated teaching tool? Perceptions of professors' teaching styles and use of humor. *College Teaching*, *52*, 14-20. https://doi.org/10.3200/ctch.52.1.14-20

Van Giffen, K. (1990). Influence of professor gender and perceived use of humor on course evaluations. *Humor*, *3*(1), 65-73. https://doi.org/10.1515/humr.1990.3.1.65

Wanzer, M. B., & Frymier, A. B. (1999). The relationship between student perceptions of instructor humor and students' reports of learning. *Communication Education*, *48*, 48-62. https://doi.org/10.1080/03634529909379152

Wanzer, M. B., Frymier, A. B., & Irwin, J. (2010). An explanation of the relationship between instructor humor and student learning: Instructional humor processing theory. *Communication Education*, *59*(1), 1-18. https://doi.org/10.1080/03634520903367238

Wanzer, M. B., Frymier, A. B., Wojtaszczyk, A. M., & Smith, T. (2006). Appropriate and inappropriate uses of humor by teachers. *Communication Education*, *55*, 178-196. https://doi.org/10.1080/03634520600566132

Wilson, J. H., Ryan, R. G., & Pugh, J. L. (2010). Professor-student rapport scale predicts student outcomes. *Teaching of Psychology*, *37*, 246-251. https://doi.org/10.1080/00986283.2010.510976

Wilson, J. H., & Taylor, K. W. (2001). Professor immediacy as behaviors associated with liking students. *Teaching of Psychology*, *28*(2), 136-138.

Yip, J. A., & Martin, R. A. (2006). Sense of humor, emotional intelligence, and social competence. *Journal of Research in Personality*, *40*, 1202-1208. https://doi.org/10.1016/j.jrp.2005.08.005

First Generation College Students' Perceptions
of an Academic Retention Program

Lisa Schelbe
Florida State University
lschelbe@fsu.edu

Martin Swanbrow Becker
Florida State University

Carmella Spinelli
Florida State University

Denesha McCray

Abstract: This qualitative study examines the perceptions of students enrolled in a campus-based program designed to promote academic success and retention of first generation college students. Method: Twenty-five undergraduate students in the program participated in focus groups and interviews to share their perceptions and experiences. Research team members conducted a thematic analysis on the focus groups and interviews transcripts. Findings: Students reported program components that contributed to their academic success and retention including support, expectations, resources, and preparation. Students also described concerns about how students' needs changed over time and how students in the program were perceived on campus.

Keywords: First generation college students; higher education; retention; college transition; academic success

First-generation college students are the first in their family to attend college–neither parent has attended college nor has been awarded a college degree (Padgett, Johnson, & Pascarella, 2012; Stebleton, Soria, & Huesman, 2014). The strengths and resiliency of first generation students cannot be overstated as many have overcome substantial obstacles to pursue higher education. In 2010, an estimated 4.5 million first generation college students were enrolled in colleges and universities in the United States (Pryor, Hurtado, DeAngelo, Blake, & Tran, 2011). With the high number of first generation students, college and university administrations increasingly have recognized that these students face different challenges and needs when compared to their peers (Pryor et al., 2011; Reid & Moore, 2008; Stebleton et al., 2014).

Once enrolled, first generation students are four times more likely than their peers to drop out at the end of their first year (Engle & Tinto, 2008). This trend is not unique to the first year experience as compared to their peers with parents who graduated from college, using national data sets, first generation college students were found to be 8.5 time more likely to drop out of college during the first four years at college (Ishitani, 2006). A study also using national data found at the end of five years in higher education, first generation students are similarly less likely to have remained in college and earned a bachelor's degree than their peers (Pascarella, Pierson, Wolniak, & Terenzini, 2004). To address this disparity, institutions have increasingly developed programs to increase the academic successes and retention of first generation college students. This study examines the perceptions of

first generation students related to their enrollment in GenOne[1], a campus-based program designed to assist first generation college students at a large southeastern university.

Literature Review

Quantitative and qualitative studies have found when compared to their peers, first generation students are more likely to come from lower socio-economic families (Choy, 2001; Soria & Stebleton, 2012; Terenzini, Springer, Yaeger, Pascarella, & Nora, 1996) and racial/ethnic minority backgrounds (Choy, 2001; McCarron & Inkelas, 2006; Stebleton et al., 2014; Terenzini et al., 1996), work more hours while in school (Barry, Hudley, Kelly, & Cho, 2009; Pascarella et al., 2004; Stebleton & Soria, 2013; Terenzini et al., 1996), study fewer hours (Pascarella et al., 2004), and have access to fewer supportive resources (Barry et al., 2009; Padgett et al., 2012; Reid & Moore, 2008; Stebleton et al., 2014). First generation college students also expect to rely on student loans more than their peers (Lee & Mueller, 2014; Somers, Woodhouse & Cofer, 2004).

First generation students begin college at a disadvantage compared to their peers in terms of social support, academic expectations, academic preparation, and access to resources (Barry et al., 2009; D'Allegro & Kerns, 2010; D'Amico & Dika, 2013; McCarron, & Inkelas, 2006; Padget et al., 2012; Pascarella et al., 2004; Reid & Moore, 2008). The transition from high school to college can be challenging for all students, however, first generation students face distinct and more challenges than their peers (Padgett et al., 2012; Reid & Moore, 2008; Schademan & Thompson, 2015; Stebleton & Soria, 2013; Stebleton et al., 2014). One challenge is lack of financial resources, which sometimes contributes to students' need for employment while in school. The time commitment of employment can create challenges for first generation students (Mamiseishvili, 2010; Stebleton & Soria, 2013) who also have the challenge of more family commitments than their peers (Lohfink & Paulson, 2005; Stebleton & Soria, 2013). All of the challenges, if unaddressed, potentially can disrupt this transition to college and place first generation students' academic success at risk.

First generation students, having little to no experience with college life, often begin college uncertain of what their new role as a college student entails (Pascarella et al., 2004; Reid & Moore, 2008; Soria & Stebleton, 2012). As compared to their peers, first generation students begin college less confident and more unfamiliar of the coursework expectations in a university environment (Barry et al., 2009; DeFreitas & Rinn, 2013; Padgett et al., 2012; Reid & Moore, 2008). Additionally, first generation students are less academically engaged (Pascalla et el, 2004; Pike & Kuh, 2005; Soria & Stebleton, 2012; Yee, 2016). First generation students tend to be unaware of the academic expectations professors hold for their students and less likely to reach out to faculty for help or assignment clarification (Soria & Stebleton, 2012; Schademan & Thompson, 2015; Yee, 2016). In addition to disconnected expectations, first generation students often begin college without the skills necessary to succeed academically when compared to their peers (Padgett et al., 2004; Reid & Moore, 2008; Stebleton & Soria, 2013). Compared to their peers, first generation students are less confident in their academic abilities (DeFreitas & Rinn, 2013; Reid & Moore, 2008), have a lower sense of self-efficacy (Inman & Mayes, 1999), have lower educational aspirations (Pike & Kuh, 2005), and received less academic preparation during high school (Atherton, 2014; Choy, 2001; Pascarella et al., 2004). For many students, parents can provide context to the challenges of rigorous coursework and developing strong study habits, but first generation students likely do not have parents or social network members who understand these requirements of college (Padgett et al., 2004; Reid & Moore, 2008; Sy, Fong, Carter, Boehme, & Alpert, 2011). In essence, first generation college students are disadvantaged as they lack the proper exposure or mentoring to prepare for college (Soria & Stebleton, 2012).

---

[1] A pseudonyms is used for the program name to maintain confidentiality.

First Generation College Students' Resilience and Protective Factors

Despite the challenges they face, first generation college students demonstrate great resilience. Research frequently is framed in a deficit model examining the individual challenges first generation students encounter (Spiegler & Bednarek, 2013). However, many first generation students succeed in college, and understanding their resilience and strengths is important (Clauss-Ehlers & Wibrowski, 2007). Studies examining the negative outcomes (i.e., dropping out) often fail to acknowledge that first generation students often overcome significant odds to even start college.

Social support, defined as a sense of connection, belonging, and relationship satisfaction, is identified as a protective factor in the retention of first generation students (Pascarella et al., 2004; Soria & Stebleton, 2012; Wang, 2012). This type of support is influential in determining the level of social and academic engagement students have with their peers, college organizations, and faculty (Barry et al., 2009, Soria & Stebleton, 2012). For first generation students, the greater the sense of social support and belonging, the less likely they are to drop out, and the more likely they are to remain in university in pursuit of graduation (Pascarella et al., 2004; Soria & Stebleton, 2012).

Unfortunately, research has also found first generation students are less likely to engage in supportive experiences such as group study, joining student organizations, or reaching out to peers and faculty for assistance (Pascarella et al., 2004; Pike & Kuh, 2005; Soria & Stebleton, 2012; Yee, 2016). This may be because first generation students who work more and have more family commitments may not have the time necessary to become engaged in the activities (Lohfink & Paulson, 2005; Mamiseishvili, 2010). While first generation students are less likely to be engaged, those who are involved in extracurricular activities derive greater benefits and outcomes from their involvement (Pascarella et al., 2004).

Campus-based Programs

Recognizing the distinct and growing needs of first generation students, institutions have developed programs geared toward improving the retention and academic success of these students. These programs can be "bridge programs" which focus on first generation students' transition to a university, supportive programs that are available to the students throughout their academic career, or a combination of these two program types. Programs that encompass initiatives aimed at developing preparatory academic skills and building strong social and academic support networks to retain first generation students increase the likelihood they will graduate (Petty, 2014; Reid & Moore, 2008). Research has found programs for first generation college students can increase retention as well as grade point averages in the first year of college (Folger, Carter, & Chase, 2004).

Building programs to mitigate the differences between first generation students and their peers is needed; however, without direct input from first generation students themselves, programs may not fully understand how to effectively meet student's needs, and they may not fully appreciate the relative impact specific components of these programs have on students. Currently few studies examine the experiences of first generation college students from their perspectives (Byrd & MacDonald, 2005; Nichols & Islas, 2015; Reid & Moore, 2008; Stebleton et al., 2014; Swanbrow Becker, Schelbe, Romano, & Spinelli, 2017; Yee, 2016). This study addresses this gap by examining students' perceptions in terms of how GenOne, a campus-based, first generation college student program developed to promote academic success and retention, impacted their academic performance. The study's research questions are 1) how do GenOne students perceive elements of the GenOne program as contributing to student retention and 2) how do GenOne students believe GenOne could increase student retention?

Method

This qualitative study, conducted in collaboration with the GenOne Director and administration, is a thematic analysis of data from focus groups and interviews to examine the perceptions of students in GenOne, a first-generation student academic retention program at a large southeastern public university of approximately 40,000 students. The retention rate for the university is approximately 72-74% and while the retention rate for GenOne students is lower at approximately 68%, this difference in retention rate appears smaller as compared to those found in other studies (Engle & Tinto, 2008; Ishitani, 2006). The goal of the GenOne program is to promote academic success and increase retention of first generation students. GenOne administration developed the research questions with the research team, who collected and analyzed the data. The Florida State University Institutional Review Board approved the study.

Description of GenOne

GenOne serves approximately 1,500 students, the vast majority of whom are African American and Latino. The mission of GenOne is to assist first generation college students' transition to and graduation from college. In this institution specific program, students are admitted to the university through the GenOne admissions process that takes into consideration the unique challenges due to educational and economic circumstances and the potential for unique contributions of individuals. The cornerstone of GenOne is the Summer Bridge Program, an intensive eight-week program that provides incoming students with orientation to the program, university, and community. During the Summer Bridge Program, staff assign each GenOne student to a small group led by an ambassador, an upper class student serving as paid peer mentor. Additionally, students connect with GenOne staff during program events and informal interactions.

The Summer Bridge Program seeks to provide experiences to help students prepare for the academic rigors of college. Students attend introductory college classes, exposing them to new, heightened academic expectations and the necessity of building new study and time management skills. Monitoring of student participation occurs through mandated study hours in a dedicated study lab that provides access to tutors and through interactions with the ambassadors.

In order to support students' social, emotional, and academic growth, students gain access to the institution's resources through shared living and programming. During the Summer Bridge Program students live in university housing and often form close bonds with their peers. They attend meetings where they are introduced to the institution's representatives and resources (e.g., counseling center, recreation facilities). They also engage in peer education programs to provide the scaffolding by which more advanced GenOne students share advice and guidance to the new students.

Following their introductory summer and throughout their first year of college, GenOne students are required to utilize study lab and log a minimum number of hours per semester starting at 10 hours, but then transitioning to between 5 and 13 hours based on their grade point average. This mandated study time seeks to help them maintain focus on academics and build effective study habits. In addition, tutors are available at the study lab to assist with English and math. Students access academic advising to assist with registering for classes and answer questions about the curriculum. Throughout the first year, students participate in monthly general assemblies where they maintain connection with GenOne students and staff, learn about events and the institution's policy, and gain access to programming to support their adjustment to college.

One of the strongest design features of the GenOne program promoted by administration is the connection with staff that is established during the summer and maintained for many throughout

their college years. GenOne staff strongly advocate for the students and can assist them in times of need, such as when students encounter financial struggles or difficulties adjusting to college.

After successfully completing the first year of college, services are reduced. During their second year of college, individual coaching is provided to GenOne students to discuss any aspects of their life and to develop plans to achieve their goals. Coaches can help students explore issues with academics, career choice, personal development, and relationships. After the second year in GenOne, fewer structured programs are offered to students. Many students, however, choose to continue their involvement with events and may continue using the tutoring and study lab.

Sample

All study participants were involved with GenOne. A research team member introduced the study during a GenOne meeting and invited students to participate in a focus group. Additionally, a GenOne staff member e-mailed students about upcoming focus group sessions, directing all interested participants to the research team to confirm attendance. Participants consented to be interviewed about their thoughts and experiences with GenOne. Faculty members and trained graduate students worked in pairs to co-facilitate the focus groups with GenOne students and individual interviews with GenOne ambassadors. Compensation was not provided to participants; however, food and drinks were offered during the focus groups.

Faculty members and trained graduate students co-facilitated four focus groups and six interviews. Focus groups lasted between 27 and 180 minutes ($\bar{x} = 71$), while interviews lasted between 19 and 72 minutes ($\bar{x} = 45$). Study participants included 25 GenOne students, four of whom were GenOne ambassadors. Saturation was reached with the sample size, which is consistent with previous research that has reached saturation with samples of as few as a dozen participants (Padget, 2017). The final sample consisted of 19 female and six male students. Of those, 17 were first year students, three were second year students, and five were upper-class students. The students were between 18 and 22 years old with the average age of 19.1 years. Students self-reported ethnic and racial identifications as follows: 14 as African American; eight as Hispanic; two as White, and one as Other.

Data Collection and Analysis

Data were collected over a four-month period in the spring of 2014 during which time the research team conducted focus groups with GenOne students and individual interviews with Ambassadors. This project collected data through focus groups with students to incorporate perspectives from different people while providing the opportunity for reaction to comments from other participants (Braun & Clark, 2013). Interviews were conducted with Ambassadors because interviews provided more time to ask follow-up questions and get more details (Braun & Clark, 2013) about the Ambassadors' experiences who not only were GenOne students, but also had experiences working with other GenOne students throughout the Summer Bridge program. All focus groups and individual interviews were audio recorded and transcribed verbatim.

The research team analyzed data in accordance with the principles of representational thematic text analysis, where themes emerge directly from the data rather than from pre-determined categories. The analysis plan followed the guidelines for thematic analysis outlined by Braun and Clark (2013), which occurs over six steps: become familiar with the data; generate initial codes; search for themes; review themes; define and name themes; and produce the report. Three team members individually read and coded each transcript to identify themes related to the research questions. Following coding, team members wrote memos reflecting on the transcript contents and any themes. The team met

regularly during the analysis process to discuss the data and through discussion reached consensus on the themes in the data.

The research team employed multiple strategies to increase the trustworthiness and rigor of the study. The study used a thematic analysis process where multiple team members coded independently and reached consensus in each step (Braun & Clarke, 2013). The diversity of participants and multiple focus group and interviews allowed for triangulation of data. The research team's diversity also contributed to the trustworthiness of the study as the team included faculty and research assistants from different disciplines, racial identities, and genders; research assistants pursuing different levels of education (i.e., bachelors, masters, and doctoral); and one research assistant who was a first generation college student. Throughout the process, the research team regularly debriefed and engaged in reflexivity.

Findings

Students endorsed several components of GenOne that they felt directly led to their academic success and retention, including *support, expectations, preparation,* and *resources. Support* encompasses the beneficial relationships of staff, faculty, mentors, and peers. *Expectations* is conceptualized as having a sense of needing to be responsible in college and planning to succeed. *Preparation* includes personal growth and skill development attributed to the programs that assisted with their transition to college. *Resources* include tools, institution connections, and financial assistance.

When asked about their ideas for how GenOne could better serve students and increase retention, students consistently emphasized their desire for GenOne to continue to provide services past the early college years. Two areas that were highlights as opportunities for improvement included *attending to the changing needs* of GenOne students over the course of their academic careers and *addressing campus perceptions* and reducing stigma faced by the first generation college students.

Support

Students most frequently attributed their connection to GenOne staff and ambassadors as the strongest factor impacting their transition to college. A common theme emerging from the data suggests that those students able to form these connections continued to access them throughout college experience. Some students found value in having developed a relationship with someone in GenOne to reach out to in times of need but also someone who may check in on them periodically. One student explained about GenOne staff, "[T]hey really believe in your dreams so whenever you have something that you need help in or something that you want to do to help you with your future, they really support you." Often while discussing support, students mentioned having a similar background or shared common interests with staff members they accessed, such as belonging to the same sorority or fraternity.

Students report that the strong personal connections developed through GenOne lead them to feeling part of the GenOne "family." One student shared, "It was really a family. It wasn't just a program." Students frequently noted that at times they felt like they may have to leave college due to various circumstances, but they persisted because they did not want to let down GenOne staff, their ambassador, and their peers. A student explained, "It was almost like [staff] were a parent figure. And it was like we don't want to let you down because you brought us in. Why would we put you to shame?" One student described GenOne as "my home away from home."

Students were clear that while GenOne contributed to their successes, support became critical to their retention during a crisis. One student shared about her previous semester:

I was failing two or three of my classes. And I was going through a lot emotionally and I was in counseling. And I think that [staff] pulling the student aside—which is what they did with me—and talking through it and giving me an action plan to execute and meeting up again to check in really helps. And that's definitely what kept me here and what's keeping other people here.

While such connection served many students, the data suggest it had a differential impact for those students not able to form solid connections with ambassadors or staff early on. In these cases, students tended to get "lost." For students who ask GenOne for help, their connections to people served them well in terms of acquiring support. Those not asking for help, however, were not tracked by GenOne; staff did not know these students needed support and consequently a subset of students who may need help were less likely to find support. An exception is a program within GenOne for former foster youth. Students in this program had regular meetings and check-ins from a dedicated GenOne staff member throughout their time at the university.

Expectations

Many GenOne students reported that as the first in their families to attend college, they lacked role models for academic success and GenOne filled this gap through providing role models and expectations for students that promoted accountability. Many ambassadors and staff had experienced similar struggles as the students and persevered, which students said was both motivating and also helped form interpersonal connections. Several participants discussed how, as students experienced hardships and considered leaving the university, staff both inspired them and reminded them about the importance of education. Students describe a culture of success where GenOne staff maintain the belief that students should remain at school despite hardships they may encounter. "They don't let you go easy," one student said referring to GenOne staff discouraging students from leaving school. Students were clear that while there was support and resources available to assist students, ultimately students were accountable for their success and staying in school. On student summed this up, "It's on you." Such clear articulation of responsibility appeared to help students focus on their goal of succeeding in college.

Students discussed how staff held them to high standards and challenged them to succeed, as is evident in one student's remark:

I know stepping out of my comfort zone was something good for me. But [GenOne] pushed me to be more vocal and meet new people and join organizations. So that is what I did. I'm grateful that they did push me to do that.

Students explained the staff's high expectations and holding them accountable often filled a gap in their lives as their family members and cultural expectations from their home communities did not consistently demand the same excellence.

Mandatory study hours and monthly general assemblies helped provide structure and clear expectations for students to focus on their academics. Many students confessed they while they had not fully developed the skills or habits in high school to manage their time in a way where they could handle the academic rigors of college, they also had not understood the expectations to do so. The standards of higher education were not clear to them prior to coming to college, and students reported feeling ill-prepared to meet the expectations to succeed. The Summer Bridge Program offered a nurturing, yet rigorous, introduction to college and helped set new expectations. The ongoing

programing and relationships with staff then reinforced the expectations.

Preparation

The Summer Bridge Program was almost synonymous with GenOne for many of the students who talked about how being in the program helped their transition to college. One student explained:

> It really, like, helped me feel comfortable. And all the information helped me build my confidence and going through the, I mean, having the [GenOne study] lab hours and stuff like that… helped me academically…it just helped me feel way more confident in myself...

> Further, by living together in university housing during the Summer Bridge Program, students were able to practice more independent living within a supportive structure. Students noted that mandatory study hours with tutors available assisted them in developing time management and study skills to help them adjust to the academic rigors of college. A student explained, "Although the university [helps] during orientation, I felt that GenOne did a great job of making [orientation] on a more of a personal level." Students repeatedly disclosed that GenOne helped prepare them in terms of personal growth and skill development for the rigors of college, which ultimately helped them succeed and stay in school.

Resources

Access to and knowledge of resources provided another aspect of GenOne that students reported as instrumental to their ability to thrive academically and socially. Summer Bridge Program and general assemblies throughout the first year exposed students to resources such as the counseling center, financial aid office, academic advising, and dean of students office, helping them with assimilation into their new environment. A student explained, "[GenOne] really did put resources out there for me and guided me in the right direction." Student reported that access to these resources helped them engage in healthy behaviors (e.g., healthy eating, exercise) and also access resources more quickly in times of need (e.g., counseling, advising). Scholarships and grants were specifically mentioned as resources that helped students remain in school.

Further, students experiencing a significant hardship, often described as financial troubles at home and academic concerns, reported GenOne staff made significant contributions to helping them resolve their issues. GenOne staff would often advocate for students across the institution and connect students with important resources, such as the Dean of Students Department and Financial Aid. A student commented that GenOne staff "moves mountains for us" when students needed help. Another explained, "They go above and beyond if you need anything." Students routinely shared that they received help when they were in crisis. One student explained,

> [staff] gave us all the tools to do it and if we wanted…we could get help. That's just the initiative of the student, they already have told us like, "Here this is how you get help. Here is where you go. like, this is where you can go to talk about it."

Students believed the connection to the resources along with the assistance from GenOne staff significantly contributed to their retention.

Opportunities for Improvement

When students were asked about how GenOne could better assist first generation college students, they most frequently suggested extending the current program offered beyond the early college years. Students expressed a desire for continued connection to the program and for the ambassadors to continue their support throughout their first year and into their second year. Similarly, students requested the programing continue throughout their college career. While most discussions about how GenOne could improve were positive, there were some concerns raised, specifically regarding how GenOne students' needs changed over time and how GenOne students were perceived across the institution.

Attend to Changing Needs. One concern expressed by students lies in the shifting needs of students as they progress through college. Following the intensive Summer Bridge Program, primary support from GenOne shifts to monthly general assemblies, mandated study hours, tutors, and academic advising over the first year. With the second year those supports generally drop off. Support in the junior and senior years was available upon request, but no formal support was in place.

GenOne provides intensive support to aid in student transition to college, especially through the Summer Bridge Program and the first year with the required study hours and general assemblies. The shift from the intensive support being provided with mandatory program to support being offered at students' initiative was jarring for some. One student described this:

> [GenOne] just kinda push you out of the nest, like they expect you to fly… I felt like they were there all the time. And then all of a sudden, they were just gone. And it was like, "Okay, now it is your responsibility."

As their college career progressed, students had to be more proactive to stay connected.

In the second year, students endorsed having more existential needs, such as finding a major, thinking about a career, managing increasing independence and responsibilities, and more generally transitioning into adulthood. Coaches were available to those students who choose to take advantage of such resources, but more structured programming directed to all students and these particular developmental needs may help the broader group, students explained. Students stressed that having ambassadors continue throughout the first and second year would be helpful; although several also acknowledged that ideally the ambassadors would be compensated for their assistance.

Students reflected that the needs for access to resources and support from GenOne continued after the first and second year. One student shared, "And I understand that there is a lot of things that go on in life, and I understand that you can't continue to keep getting babied, but to some extent we're going to need that information." There was the general consensus that throughout college, students would benefit from GenOne programs. One student summed up the general consensus that there was a need to implement a "more aggressive approach to getting students to be active [in GenOne] after their freshman and sophomore year" because "after your freshman and sophomore year, you do have a different mindset, not only about school but about a lot of things."

In the final year of college, students expressed needing help with managing their career search and deciding whether to attend graduate school. Students described how they wished GenOne would attend to the changes needs of students as they progress through their studies. Several students in different focus groups mentioned ideally GenOne would continue for students who pursued advanced degrees in graduate schools.

Students expressed a desire for continued support throughout their time in college and GenOne to address the changing needs of students. However, they quickly emphasized that they did not want to feel "babied" and were conflicted about making any programs mandatory. Students expressed wanting to have freedom in choosing to be involved in GenOne throughout college, yet

some pointed out the issue that that optional programs may not be fully utilized. A student explained, "I feel like making [programs for juniors and seniors] mandatory would help…I know for certain people, they are like, 'oh, it's not mandatory, so I am not showing up.'" Indeed, one student reflected:

> Once you hit the junior year [general assemblies and study hours] aren't mandatory anymore, but they are still helpful if you chose to go… It's not up to [GenOne] if you want to go or not, but they are giving you the option. So they are putting that resource out there. It's your fault if you take advantage of it or not.

This tension between wanting support, yet not having mandatory programming was throughout students' responses about how GenOne could better serve students.

Addressing Campus Perception. Students endorsed a deep sense of pride in being part of GenOne. One student stated this strongly: "I may not always bleed [the university's colors], but if you cut me I will bleed [GenOne]. That is just what it is." Along with the sense of pride and identity was an acknowledgement of a stigma at the institution regarding GenOne students, with some saying that they felt GenOne students were perceived as not belonging at the institution due to the alternative admission criteria and as being "babied" with the extra resources and support available to them.

The alternative admission criteria used by GenOne meant that on average high school grade point averages and standardized test scores of GenOne students were lower than that of the institutions' general student body. Students described how others believed that the alternative admission criteria translated to GenOne students not seen as being legitimate students on campus. Several students described overhearing conversations about GenOne students where other students called GenOne students "dumb" and said "[GenOne students] are not supposed to be here." One student described defending the alternative criteria of GenOne and responded, "There's some bright people [at GenOne]. It's just we didn't get in the same way you did."

GenOne students were adamant that they belonged at the institution and defended their admission due to the circumstance related to their being first generation students. One student shared, "I didn't get in [to the institution] on my SAT or ACT scores because I had a hard time paying for it." Despite their belief that they belonged at the institution, GenOne students found the stigma and perceptions of GenOne taxing. One student explained:

> It's very hard being a [GenOne] student because not only are people like… "How'd y'all get into that program?", "Oh so y'all got in the easy way?", and stuff like that…and a lot of those people are saying "Oh, you only got into [university] because of [GenOne]."

GenOne students emphasized that consistently grade point averages of GenOne students were equal to those of their peers at the university. Likewise, students mentioned GenOne students disproportionately were student leaders in camps groups.

Stressing the majority of GenOne students were African American and Latino, students across multiple focus groups emphasized the perception of not belonging at the institution was also racially and ethnically influenced. One student explained, "It's a very crude message that I hear from people…people have this perception that [GenOne]'s an ethnic-based like group organization rather than just being something that helps with people that didn't have the chance." Many students described there was a negative stereotype about GenOne only being for African America students who were not qualified to be at the institution. Specifically, a few students emphasized that the university was a predominantly white institution and explained that they experienced racism on campus. Some of the racism faced was related to the stigma of GenOne and doubting the academic

abilities of people of color. One student explained:

> The connotation of [GenOne] is that it is for African American students and that is the only way they can get into college…people see it as [GenOne] is the only way Black people can get into this institution and that's not the only way it is. But [GenOne] helps our voice be recognized.

This student emphasized that while the stigma of GenOne impacted him, overall the program also assisted him and other students of color. Another student shared a similar point, "When I first got here, the race and discrimination and the things that were said, they just got to me…[GenOne] taught you how to deal with it and cope with it." Thus, despite the stigma and discrimination, students frequently credited GenOne for helping address the problems.

Another part of the stigma GenOne students experienced was connected to the support GenOne provided. Several students from multiple focus groups explained a history of nurturing the GenOne students led to the informal label of "[GenOne] Babies," a reference to a belief that GenOne students are overly nurtured and provided with resources not available to other students. Connected to this stereotype was that the students would not be able to attend the institution without extensive support as they did not truly belong there. Students recommended marketing the program to the broader institution to help the university community better understand the value of GenOne students and work to establish a more positive image to the program. One student emphasized, "[GenOne] is doing all they can to put us out there, but the university could give us the publicity that we need." Students described a desire for publicizing facts about GenOne students' grade point averages being equal to their peers at the institution and their involvement as student leaders. Likewise, GenOne students wanted their successes to be highlighted as examples of how first generation students can thrive. It was important to students that the stereotype of being "babied" was dismantled. Other students suggested increased attention to the balance between nurturing students and challenging them may help such programs address concerns of stigma and facilitate student growth.

Discussion

The insights provided by GenOne students informs thinking about how programs may better serve first generation college students. As previous studies of first generation college students have found, GenOne students faced obstacles to remaining in school. Social support, which studies have found to be a protective factor for first generation students to remain in school (Pascarella et al., 2004; Pike & Kuh, 2005; Soria & Stebleton, 2012; Wang, 2012), that students received through GenOne was invaluable and they perceived it to be central to their success in education. Connecting to their peers provided a sense of family, as Jehangir (2009) found. Similar to findings in other studies (e.g., Pascarella et al., 2004; Reid & Moore, 2008; Soria & Stebleton, 2012), GenOne students reported guidance and assistance in transitioning to college and understanding the culture and expectations for college students assisted them academically.

Research consistently has found that first generation students face use unique challenges during the transition to college (Padgett et al, 2012;Reid & Morre, 2008; Schademan & Thompson, 2015; Stebleton & Soria, 2013; Stebleton et al., 2014) and campus support programs can assist with the transition (e.g., Folger, Carter, & Chase, 2004). The transition to college was not the only time that GenOne students identified as needing support. They emphasized their needs changed throughout their college careers and it would be beneficial to have ongoing support and different resources as their needs changed. For example, approaching graduation, students may benefit from assistance in job searching or applying to graduate school.

The GenOne students' desire for continued programs and support was somewhat internally inconsistent. They were adamant that while they wanted additional support, they did not want programs to be mandatory. This is despite their recognition that upper class students could still participate in the general assemblies, study hours, and other programs after their sophomore year, yet few did. Students explained this suggested students likely would only would participate if programs were mandatory.

Students expressed concern about the stigma they faced in terms of the alternative admission criteria and being perceived as being "babied" by GenOne. They did not want their peers to see them as unworthy or believe that GenOne students were only successful because of the extra support they received. The frustration with stigma was inherently connected to racism and discrimination as the majority of GenOne students were African American and Latino.

The concerns GenOne students raised about stigma and the discrimination they faced and some of the internal inconsistencies and ambiguity could potentially be understood better considering the impact stereotypes (Aronson & Inzlicht, 2004) and stigma (Brown & Lee, 2005) have been found to have on academic performance. Stereotype vulnerability, the tendency to be influenced by negative stereotypes about one's social group(s) can impact students' academic self-knowledge and performance (Aronson & Inzlicht, 2004). Likewise, in one study, compared to white and Asian students, Black and Hispanic students experiencing stigma with high levels of consciousness about stigma have been found to have lower grade point averages although there was no difference in grade point averages found in Black and Hispanic students those with low levels of consciousness about stigma (Brown & Lee, 2005). GenOne students intuitively seemed to be aware that the stigma and stereotypes they faced potentially negatively impacted them.

Implications

These findings support a call for institutions to consider providing resources to further academic retention throughout the college years and to tailor those resources to the emerging needs of first generation students as they shift from transitioning from high school to college to preparing for graduation. Admission of first generation college students and getting them to campus is only the beginning. Developing the presence of broad support for these students and opportunities for strong relationships with faculty, staff, mentors, and other students is paramount. Students noted that the personalized attention they received helped them feel like part of a family, but to provide such focus programs must have sufficient staff to attend to student's needs and also track students who do not proactively seek support. Relatedly, staff must balance the provision of support with a focus on encouraging student growth and resiliency by providing enough, but not too much help. Students must also have a clear understanding of the expectations for them as college students as they may not have had family members and friends who can provide appropriate guidance. Preparing first generation colleges students to be successful can be accomplished through fostering their personal growth and skill development. Additionally, programing can focus on helping first generation students' transition to institutions through connecting them with resources at the institution and in the community as well as financial assistance.

When designing and implementing programming for first generation students, their unique and changing needs as well as the larger institutional context should be taken into consideration. The needs of first generation students who are transitioning to an institution are markedly different than those of first generation students who are later in their studies or who are planning to graduate. Throughout their academic careers first generation students benefit from support and services, yet it is important to tailor programs to the students' specific needs. While focusing first generation students' transition to campus is central to retention, programs should meet their needs later in their studies

including major selection, career planning, and professional development. The needs of students change throughout their college careers and programs need to ensure that upper class students continue to be supported and have access to resources that will help them graduate and transition post-graduation.

As first generations students are disproportionally from racial and ethnic minority backgrounds, programs should take into consideration the stigma and racism that students may face. This is especially the case as predominately white institutions. There may be campus perceptions and negative stereotypes that need to be addressed at an institutional level to ensure that the environment is conducive for first generation students. Institutions should seek to ensure that all students on campus can complete their studies in an environment free from stigma and discrimination. Reducing these stressors and barriers can contribute to students' successes.

This study is limited by the small convenience sample, which was interviewed at only one time point. Nevertheless, the study makes a significant contribution to the literature by examining the perceptions and experiences of first generation college students receiving services. In particular, this study highlights components of academic retention programs that students find most impactful to further their development and help them thrive in college. Additionally, it outlines potential areas which programs can address to improve serving to first generation college students. Future research may wish to examine not only retention, but a holistic view of the experiences of first generation college students in programs such as GenOne. It may be useful to consider differences in race and gender as well as students' experiences before college. Similarly, it may be important to study stigma and experiences of discrimination that first generation students face, especially students of color. Future research could also follow first generation college students after graduation to determine long-term outcomes stemming from academic retention programs such as this.

Conclusion

Overwhelmingly, the students attributed their success and retention to the support, expectations, preparation, and resources that they received through the GenOne program.
Enrolling first generation students is only the first steps to helping them succeed. Institutions must focus on supportive programing for first generation student throughout their time in college, building a sense of community, providing mentoring, and helping students develop skills to meet their higher expectations. Through understanding the experiences of first generation college students, universities can promote college retention and also holistic personal and professional development.

References

Aronson, J., & Inzlicht, M. (2004). The ups and downs of attributional ambiguity: Stereotype vulnerability and the academic self-knowledge of African American college students. *Psychological science, 15*, 829-836.

Atherton, M. C. (2014). Academic preparedness of first-generation college students: Different perspectives. *Journal of College Student Development*, *55*, 824-829. https://doi.org/10.1353/csd.2014.0081

Barry, L. M., Hudley, C., Kelly, M., & Cho, S. (2009). Differences in self-reported disclosure of college experiences by first-generation college student status. *Adolescence, 44*, 55-68.

Braun, V. & Clarke, V. (2013). Successful qualitative research: A practical guide for beginners. Thousand Oaks: Sage.

Brown, R. P., & Lee, M. N. (2005). Stigma consciousness and the race gap in college academic achievement. *Self and Identity, 4*, 149-157.

Byrd, K. L., & MacDonald, G. (2005). Defining college readiness from the inside out: First-generation college student perspectives. *Community College Review*, *33*, 22-37. https://doi.org/10.1177/009155210503300102

Choy, S. (2001). Students whose parents did not go to college: Postsecondary access, persistence, and attainment. In: National Center for Education Statistics, editor. The condition of education (pp. xviii-xliii; NCES publication no. 2001-072). Washington, DC: U.S. Government Printing Office.

Clauss-Ehlers, C. S., & Wibrowski, C. R. (2007). Building educational resilience and social support: The effects of the educational opportunity fund program among first-and second-generation college students. *Journal of College Student Development*, *48*, 574-584. https://doi.org/10.1353/csd.2007.0051

D'Allegro, M. L., & Kerns, S. (2010). Is there such a thing as too much of a good thing when it comes to education? Reexamining first generation student success. *Journal of College Student Retention: Research, Theory & Practice, 12*, 293-317. https://doi.org/10.2190/CS.12.3.c

D'Amico, M. M., & Dika, S. L. (2013). Using data known at the time of admission to predict first-generation college student success. *Journal of College Student Retention: Research, Theory & Practice*, *15*, 173-192. https://doi.org/10.2190/CS.15.2.c

DeFreitas, S. C., & Rinn, A. (2013). Academic achievement in first generation college students: The role of academic self-concept. *Journal of the Scholarship of Teaching and Learning*, *13*(1), 57-67.

Engle, J., & Tinto, V. (2008). Moving beyond access: College success for low-income, first-generation students. Washington, D.C.: The Pell Institute.

Folger, W. A., Carter, J. A., & Chase, P. B. (2004). Supporting first generation college freshman with small group intervention. *College Student Journal, 38*, 472-476.

Inman, W.E., & Mayes. L. (1999). The importance of being first: Unique characteristics of first generation community college students. *Community College Review, 26*, 3-22. https://doi.org/10.1177/009155219902600402

Ishitani, T. T. (2006). Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education*, *77*, 861-885. https://doi.org/10.1353/jhe.2006.0042

Jehangir, R. R. (2009). Cultivating voice: First-generation students seek full academic citizenship in multicultural learning communities. *Innovative Higher Education*, *34*, 33-49. https://doi.org/10.1007/s10755-008-9089-5

Lee, J., & Mueller, J. A. (2014). Student loan debt literacy: A comparison of first-generation and

continuing-generation college students. *Journal of College Student Development*, *55*, 714-719. https://doi.org/10.1353/csd.2014.0074

Lohfink, M. M., & Paulsen, M. B. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development*, *46*, 409-428. https://doi.org/10.1353/csd.2005.0040

Mamiseishvili, K. (2010). Effects of employment on persistence of low-income, first-generation college students. *College Student Affairs Journal, 29*, 65-74.

McCarron, G.P., & Inkelas, K.K. (2006). The gap between educational aspirations and attainment for first-generation college students and the role of parental involvement. *Journal of College Student Development, 47*, 534-549. https://doi.org/10.1353/csd.2006.0059

Nichols, L., & Islas, Á. (2016). Pushing and pulling emerging adults through college: College generational status and the influence of parents and others in the first year. *Journal of Adolescent Research*, *31*, 59-95. https://doi.org/10.1177/0743558415586255

Padgett, D. K. (2017). *Qualitative methods in social work research* (3th ed.). New York: Sage.

Padgett, R.D., Johnson, M.P., & Pascarella, E.T. (2012). First-generation undergraduate students and the impacts of the first year of college: Additional evidence. *Journal of College Student Development, 53*, 243-266. https://doi.org/10.1353/csd.2012.0032

Pascarella, E.T., Pierson, C.T., Wolniak, G.C., & Terenzini, P.T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *The Journal of Higher Education, 75*, 249-284. https://doi.org/10.1353/jhe.2004.0016

Petty, T. (2014). Motivating first-generation students to academic success and college completion. *College Student Journal, 48*, 257-264.

Pike, G. R., & Kuh, G. D. (2005). First-and second-generation college students: A comparison of their engagement and intellectual development. *The Journal of Higher Education*, *76*, 276-300. https://doi.org/10.1080/00221546.2005.11772283

Pryor, J. H., Hurtando, S., DeAngelo, L., Blake, L .P., & Tran, S. (2011). *The American freshman: National norms fall 2010*. Los Angeles: Higher Education Research Institute, UCLA.

Reid, M.J., & Moore, J.L. (2008). College readiness and academic preparation for postsecondary education: Oral histories of first-generation urban college students. *Urban Education, 43*, 240-261. https://doi.org/10.1177/0042085907312346

Schademan, A. R., & Thompson, M. R. (2015). Are college faculty and first-generation, low-income students ready for each other?. *Journal of College Student Retention: Research, Theory & Practice*. https://doi.org/10.1177/1521025115584748

Somers, P., Woodhouse, S. R., & Cofer Sr, J. E. (2004). Pushing the boulder uphill: The persistence of first-generation college students. *NASPA Journal*, *41*(3), 418-435. https://doi.org/10.2202/0027-6014.1353

Soria, K.M., & Stebleton, M.J. (2012). First-generation students' academic engagement and retention. *Teaching in Higher Education, 17*, 673-685. https://doi.org/10.1080/13562517.2012.666735

Spiegler, T. & Bednarek, A. (2013). First-generation students: what we ask, what we know and what it means: an international review of the state of research. *International Studies in Sociology of Education, 23*, 318-337. https://doi.org/10.1080/09620214.2013.815441

Stebleton, M., & Soria, K. (2013). Breaking down barriers: Academic obstacles of first-generation students at research universities. *Learning Assistance Review,17, 7*-19.

Stebleton, M.J., Soria, K.M., & Huesman, R.L. (2014). First-generation students' sense of belonging, mental health, and use of counseling services at public research universities. *Journal of College Counseling, 17*, 6-20. https://doi.org/10.1002/j.2161-1882.2014.00044.x

Sy, S. R., Fong, K., Carter, R., Boehme, J., & Alpert, A. (2011). Parent support and stress among first-generation and continuing-generation female students during the transition to

college. *Journal of College Student Retention: Research, Theory & Practice, 13*, 383-398. https://doi.org/10.2190/CS.13.3.g

Swanbrow Becker, M., Schelbe, L., Romano, K., & Spinelli, C. (2017) Promoting first-generation college students' mental wellbeing: student perceptions of an academic enrichment program. *Journal of College Student Development, 58,* 1166-1183.

Terenzini, P. T., Springer, L., Yaeger, P. M., Pascarella, E. T., & Nora, A. (1996). First-generation college students: Characteristics, experiences, and cognitive development. *Research in Higher Education*, *37*, 1-22. https://doi.org/10.1007/BF01680039

Wang, T.R. (2012). Understanding the memorable messages first-generation college students receive from on-campus mentors. *Communication Education, 61*, 335-357. https://doi.org/10.1080/03634523.2012.691978

Yee, A. (2016). The unwritten rules of engagement: Social class differences in undergraduates' academic strategies. *The Journal of Higher Education*, *87*, 831-858.

Not All Flipped Classes are the Same:
Using Learning Science to Design Flipped Classrooms

Alyssa P. Lawson
University of California, Santa Barbara
a_lawson@ucsb.edu

Caylor R. Davis
California State University, Los Angeles
cdavis38@calstatela.edu

Ji Y. Son
California State University, Los Angeles
json2@calstatela.edu

Abstract: The flipped classroom has recently become a popular method used in both higher education and K-12 classrooms, yet research has not consistently demonstrated clear benefits of flipping a classroom. Also, any benefits seen might not be from the flipped design itself, but instead from the individual aspects of a flipped classroom (e.g., more active learning in class, more feedback on homework). This suggests that research focused on how to develop the activities and components of the flipped classroom, instead of simply flipping the traditional in-class and out-of-class activities, is critical to flipped pedagogy. These activities (both in and out of the classroom) should be designed based on a theory of learning. We draw upon the Practicing-Connections (PC) Hypothesis, supported by contemporary theories and research in the cognitive sciences, to design instructional activities to promote learning by having students practice making connections between concepts and situations. This paper examines a flipped class based on a theory-driven design versus a basic flipped classroom. The results of this work offer suggestions as to what dimensions of flipping may be important and how to design and evaluate flipped classrooms based on theories of learning.

Keywords: flipped classroom, active learning, learning theory, instructional design, transfer

The underlying assumption of traditional lectures is this: the delivery of content by a knowledgeable speaker will impact all levels of students' learning including remembering, understanding, applying, analyzing, evaluating, and creating. A recent meta-analysis of STEM courses revealed that the traditional lecture is, on many counts, a poorer pedagogical method compared to active learning strategies (Freeman et al., 2014). Although there is a great range of what is considered "active learning," there is rapidly growing interest in the flipped classroom (or inverted classroom). Although a basic definition of the flipped classroom is one where the activities traditionally conducted in class (i.e., lecture) and out of class (i.e., homework) swap places (e.g., Lage, Platt, & Treglia, 2000), there are cogent arguments that the flipped model's effectiveness comes from being able to utilize class time for active learning that engages higher order cognitive processes (Bergmann & Sams, 2012; Burke & Fedorek, 2017). The theory and study presented here is an effort to move past efficacy tests of the flipped classroom and towards research that is intended to help instructors *improve* the implementation of the flipped model. We will start with a review of these efficacy tests, then develop a theory of implementation based on research on learning and cognition.

Effectiveness of the Flipped Classroom

Past research on how effective the flipped classroom is for student achievement has shown mixed results. A popular method of testing the flipped model design is to compare achievement in a flipped classroom to prior versions of the same course that used more traditional pedagogies (Baepler, Walker, & Driessen, 2014; Galway, Corbett, Takaro, Tairyan, & Frank, 2014; He, Swenson, & Lents, 2012; Hill & Nelson, 2011; Wilson, 2013). Most, but not all, of these types of studies show that flipping a classroom helps students perform better on tests in a variety of domains (e.g., math: Wilson, 2013; chemistry: Baepler et al., 2014; He et al., 2012). Some of these studies show no significant differences between the two methods of teaching on student success. For example, in learning about ecosystems, students in a flipped classroom design had similar exam results to past students who had been taught in more traditional ways (Hill & Nelson, 2011). Graduate students in a flipped public health class showed no significant improvement when compared to prior students in a standard classroom (Galway et al., 2014).

Studies using more controlled methods have also been inconsistent in demonstrating the effectiveness of flipped classrooms. A study compared two separate sections of students enrolled in the same course in the same time period during a one week period; one section was designed in a flipped format for the instruction week and the other was largely traditional during the same week (Deslauriers, Schelew, & Wieman, 2011). Students in the flipped section performed significantly better on questions related to topics taught during the flipped section than students in the traditional lecture. However, other studies, using similar experimental designs, found no differences between flipped classrooms and traditional classrooms (Jensen, Kummer & Godoy, 2014; Yong, Levy, & Lape, 2015). These studies demonstrate the difficulties inherent in bridging research and instructional practice: most instructors are interested in how well students learn material that is complex, and mastered over the course of many weeks, but this prolonged study is often difficult to examine in brief, highly controlled experimental designs. Even when these studies are conducted, often with great difficulty and commitment, the results are difficult to translate back into instructional practice.

Despite the conflicting evidence on how flipped classrooms affect student performance and learning (e.g., test scores), there have been more consistent results concerning student perceptions of the flipped classroom design and student engagement. Many studies report students having positive experiences with course materials and positive attitudes towards flipped classrooms in general (Galway et al., 2014; Gilboy, Heinerich, & Pazzaglia, 2015; He et al., 2012; Hill & Nelson, 2011; Nouri, 2016; Wilson, 2013). Students' overall engagement and attendance increased when introduced to the flipped classroom design (Baepler et al., 2014; Deslauriers et al., 2011). Student reactions tended to be positive towards the flipped method, although there are documented cases of negative attitudes towards the use of technology (Jensen et al, 2014). However, this mostly positive perception may be in part a bias in the published research; researchers and instructors who have made the effort to flip a course may be less likely to go through the rigors of publication to report students' negative flipped experiences.

There is also evidence that the different components of the flipped classroom are independently beneficial to students. Meta-analyses reveal the benefits of using videos and online materials not in the context of a flipped classroom (Cohen, Ebeling, & Kulik,1981; Means, Toyama, Murphy, Bakia, & Jones, 2009). In the majority of studies investigated by Cohen et al. (1981), instruction that included visually-based instruction (e.g., videos) resulted in better student achievement than traditional classrooms. Videos that involve interactive participation by participants - including being able to stop, move forward, and move backwards - allowed students to gain more information than videos that moved from start to finish with no interaction (Schwan & Riempp, 2004) or from textbooks with the same information (Merkt, Weigand, Heier, & Schwan 2011). Interactive activities,

the key face-to-face component in a flipped classroom, have also shown to benefit learning independent of flipped classroom designs (Schwan & Riempp, 2004; Merkt, et al., 2011).

As more researchers and instructors get involved in designing flipped classrooms, it is clear that the efficacy studies that examine whether the flipped model has any effect over traditional methods has little direct impact on how we design instruction and course structures. The best these studies do is convince instructors to try out a flipped design. As technology improves and the limitations of lectures are documented, the scholarship of teaching and learning in flipped classrooms needs a shift in research focus: from "should we flip the classroom?" to "*how* should we flip the classroom?" What aspects of a flipped classroom are important? What are critical components of the design? To begin to get at the *how* question, we must move towards comparisons of flipped classrooms to other versions of flipped classrooms. The research we report here examines how two different implementations of a flipped statistics class (taught by the same instructor) can lead to markedly different effects. Our results indicated that the focus on flipping itself may be misplaced because the details of *how* a class is flipped are vitally important.

Using Learning Science to Design the Flipped Classroom

Often the design of flipped classrooms is heavily influenced by available technologies (e.g., video lectures) and forms (e.g., problem-based or team-based learning). The question of *how* a flipped class should be designed should be driven by theories of how learning works rather than how flipped courses are typically implemented. As psychologists, we propose developing design principles for flipping courses based on research in the cognitive sciences.

As defined by Lage et al. (2000), a basic flipped class simply inverts lecture and homework such that the lectures are moved out of the classroom and textbook problems (typically assigned as homework) comprise the in-class activities. The great opportunity provided by the flipped format is that face-to-face time does not have to be spent on traditional lectures delivering content. Face-to-face time can instead be used for time intensive, higher-order cognitive processing, such as creating conceptual connections, evaluating the appropriateness of concepts, analyzing the similarities and differences of various approaches, and applying concepts to novel circumstances. As has been argued by several flipped researchers and practitioners (Bergmann & Sams, 2012; Burke & Fedorek, 2017), this feature, the increase in actively *practicing connections and high level thinking*, is key to the flipped model. In the next section, we briefly summarize the latest theories and evidence in the cognitive sciences to develop this design hypothesis. Later, we present an in-depth look at a single lesson to contrast a basic flipped class with a Practicing-Connections (PC) flipped class designed to maximize the use of in class time for higher order cognitive practices. Finally, we report the results of a quasi-experimental study to examine the PC hypothesis.

Developing the Practicing-Connections (PC) Hypothesis

Although most educators commonly believe that procedures need to be practiced to achieve fluency, practice is often not considered part of acquiring concepts and conceptual skills such as thinking, reasoning, and analyzing. Most often, a lecture is intended to communicate a well thought out sequence (e.g., a proof or a narrative about scientific discovery) and hearing/comprehending the lecture is considered equivalent to thinking about the concepts. However, modern theories of cognitive science (e.g., Damasio, 1989; Barsalou, 1999; Glenberg, 1997), and in particular those associated with the "embodied cognition" perspective, posit that concepts are actually more like processes where mental representations are dynamically tailored to the current needs of some situation

(Barsalou, 2003). If concepts are actually more like skills, then in order to "learn a concept" we would need to practice tailoring mental representations to fit different situations appropriately. This is the practicing connections (PC) hypothesis: the act of connecting concepts to appropriate structures, situations, and contexts *is* the process of learning (Son, Ramos, DeWolf, Loftus, & Stigler, 2018). To flip a course using this design hypothesis, we would use both in- and out-of-class activities to help students engage in practicing connections.

From research on expertise, we know that experts often have a lot of practice, (see Chase & Simon's 10-years-of-training observation, 1973) but also their practice is varied in a particular way. The development of expertise seems to include practice that is highly responsive to outcomes, errors, and violated expectations (Chase & Ericsson, 1981; VanLehn, 1989). In these studies, people who actively engage in trying different strategies and improving their previous iterations seem to develop more skill and expertise. Deliberate practice is defined as repeated exercises where the goal is improvement, (unlike a test where the goal is achievement or repeated practice where the goal is fluency) and these exercises must be varied, more complex over time, and should not result in early automatization (Ericsson, Krampe, & Tesch-Romer, 1993). If the goal of a classroom is to produce expertise in thinking, then any "practice" should be deliberate practice. In order to practice making connections in this deliberate way, students should engage in specific tasks invented to get feedback on connections, overcome conceptual errors, and tailor concepts for multiple contexts. The practice of making connections cannot mean just making the same connections repeatedly but this activity should vary and increase in complexity.

If we want deliberate practice in making connections to be the main focus of in-class time, we must change our instructional practices to be aligned with this goal. Unfortunately, there are several barriers that make this difficult. We will focus on barriers that have been researched in context of math learning, but these limitations are often present in other domains as well. One major barrier is a maladaptive belief on the part of students, that learning math is about memorizing steps and practicing procedures (Garofalo, 1989; Stigler, Givvin, & Thompson, 2010; Schoenfeld, 1989), not about sense-making and reasoning. In response to the question, "What does it mean to be good at math?" community college students commonly gave responses such as, "In math, sometimes you have to just accept that that's the way it is and there's no reason behind it" (Stigler et al., 2010). Perhaps in part because students expect it, US teachers for the most part teach math as a set of isolated procedures that need to be memorized and practiced (Stigler & Hiebert, 1999). A contributing factor might be that instructors themselves may implicitly think of learning math in a manner similar to students. Many US teachers cannot explicitly explain connections, for example, between fraction division and the contexts in which such an operation would be useful (Ma, 1999).

Beyond just beliefs about mathematics, maladaptive beliefs about the nature of intelligence may also pose barriers (Blackwell, Trzesniewski, & Dweck, 2007; Dweck, 2007). The mindset students hold about their intelligence and/or math ability can affect their ability to perform well while learning and prevent them from absorbing new information. A student who believes that intelligence, generally, or math ability, in particular, are fixed, and therefore unchangeable, might be less inclined to engage in a process of improving their mathematical thinking.

But beliefs aren't the only barriers. Thinking and connecting are processes that require a lot of cognitive resources. In other words, thinking is hard. Research has demonstrated that people who are anxious (Beilock & Carr, 2005), impoverished (Mani, Mullainathan, Shafir, & Zhao, 2013), or otherwise saddled with a cognitive load (Sweller, 1988) are less able to engage in thinking. And to close out this less-than-exhaustive list, a final barrier to practicing thinking is that it takes up a lot of time. In order to make room in the classroom to practice thinking, some learning must be off-loaded to out-of-class time in order to "cover" the required curriculum.

The PC hypothesis is borne from theories of cognition that emphasize the importance of practicing connections and considers the cognitive barriers to such practice to drive the design of a flipped classroom. A basic flipped statistics class would simply move the lecture out of class (e.g., through video lectures) and bring textbook practice problems into class time. In contrast, a PC flipped class attempts to design activities both in and out of the classroom to facilitate the deliberate practice of making connections and address barriers to such thinking.

Based on research in cognitive science (for more details on the specific studies and theoretical approaches that have informed this hypothesis see Fries, Son, & Stigler, in preparation), the PC hypothesis suggests that we must create three types of learning opportunities for students: (1) making explicit connections between concepts, (2) productive struggle (Kapur, 2008) in making those connections, and (3) deliberate practice with connections where the tasks are varied and increasing in complexity over time. In a PC flipped class, instructors and instructional designers must identify the explicit connections for students to make between concepts and representations, contexts, and core conceptual structures rather than come up with a set of disparate learning outcomes. The design of activities (both in and out of the classroom) should scaffold and promote making those connections rather than those connections being told to students. Finally, the connections must be practiced in increasingly difficult and varied circumstances rather than being made only once or practiced in a way that promotes fluency.

Designing a Class around a PC Hypothesis

So how did we translate these three learning opportunities into a design of an actual statistics class? First, we made learning explicit connections the central goal of the class. All of the in-class activities and exams were re-designed to be authentic research situations where students would have to *practice* adapting concepts to fit the particular situation at hand. The out-of-class materials were designed to provide pre-training (Mayer, 2001) of the individual concepts that would be connected through the in-class activities. This was achieved by having students watch online videos while filling out guided notes that went along with the videos (all guided notes and links to the corresponding videos are available here: http://bit.ly/statsteach).

To provide opportunities for productive struggle, we utilized instructional technologies such as clickers and writing worksheets to give students low stakes opportunities to make connections and receive feedback. We also implemented cycles of productive struggle into the instructional routine; at the beginning of class, students could make connections from a more limited set of choices, but at the end of the class session students were challenged to produce their own explanations with less scaffolding.

Deliberate practice was implemented by designing connecting activities in a progressive sequence to be more complex and varied in systematic ways over time. One way this was implemented was that the in class activities were cumulative, such that every statistical technique students previously learned was folded into the next week's activities. For instance, if students learned about *t*-tests in one week, then in the lessons on new information that followed, students would still have to consider whether a *t*-test would be appropriate for some situations. This is unlike many math and statistics courses wherein, if the focus of the week is regression, students only practice creating regression models. In the PC flipped course, students constantly had to adapt their prior learning as they were tasked with selecting the appropriate analysis from their newly acquired repertoire of statistical tests.

At the same time, the class was designed to remove or address barriers that prevent students from focusing on making connections. Activities in class explicitly addressed cultural beliefs about mathematics and their mindsets about the nature of intelligence. We also off-loaded peripheral course

material (that must be "covered" by the class) to outside class time to give students time and space in class to engage in making connections. Additionally, we provided social supports to encourage making connections in low-stakes situations so that students could learn what it felt like to be wrong and to see the value of testing out an idea.

As with all translation of research into practice, the details matter immensely. Even though the design philosophy is based on research and theory, every lesson made many design decisions based on a number of assumptions. Because we know these details matter to instructors, we have made all in-class activities available (as worksheets) at this website: bit.ly/statsteach. In the following section, we will provide a detailed look into the design of one PC flipped lesson sequence to see how a theory of implementation might guide actual implementation. To start, we will provide a contrast: the same lesson taught in the basic flipped class and the PC class.

Basic versus Practicing-Connections Lesson Design

The aim of this work is to advance a theory of implementation based on research. We enter into this as both researchers, deeply versed in the laboratory research of our field (cognitive sciences), and instructors, grappling with the problem of coordinating teaching actions over time and across a whole class. Even though we will delve deeply into just one lesson sequence (out-of-class and in-class) in each of the flipped models (basic and PC) here, this in-depth example is meant to illustrate the design demands on instructors trying to teach productively with the flipped model.

The lesson sequence described here will tackle something covered in all statistics courses: the arithmetic mean. In both the basic and PC flipped lesson sequence, students were assigned to watch the same lecture videos that week (a 15-minute lecture on visualizing distributions in plots, http://bit.ly/lecturevideo_visualizations, a 32-minute lecture on central tendency in distributions, http://bit.ly/lecturevideo_centraltendency, and a 31-minute lecture on measures of variability, http://bit.ly/lecturevideo_variability). In the basic flipped course, there was a 10-minute open note quiz at the beginning of each class meant to encourage students to watch the videos and take notes (the quiz is provided in Appendix A). Most of the questions were drawn from verbatim statements made in the videos.

In the PC flipped course, instead of merely encouraging compliance, that is, watching the assigned videos, we conceptualized the out-of-class activities as "pre-training" (Mayer, 2001), where students would gain initial familiarity with the names and characteristics of key concepts. That would facilitate the in-class focus on creating connections between these concepts. The PC flipped class replaced the quiz with guided notes, which are outlines of the lecture with blank spaces for students to fill in with presented content, and have been shown to facilitate students' recall of complex information (Austin, Lee, Thibeault, Carr & Bailey, 2002; Neef, McCord, & Ferreri, 2006). These guided notes were due at the start of class and checked for completion, and then utilized by the students as a reference during in-class activities (The associated notes for this lesson are numbered 03a, 03c, and 04a and are available here: http://bit.ly/StatsTeach_Notes).

Note that both the basic and PC flipped designs require a commitment to active learning during the in-class activities. In the basic flipped course, to learn about the arithmetic mean, students completed chapter problems from a well-developed textbook from accomplished statistics education researchers (Watkins, Scheaffer, & Cobb, 2nd edition). Some of these problems were modified so that students could practice data analysis using statistical software. The full set of questions are provided in Appendix B.

If we zoom in further at just the parts of the lesson aimed at teaching students about the arithmetic mean, we can see that students were asked a variety of calculation oriented questions from

calculating the mean and median from small sets of numbers (e.g., problem 1a: 1, 2, 3; problem 1d: 1, 2, 297), adding a new data point and adjusting the mean, and confirming that the residuals from the mean $(x_i - \bar{x})$ add up to 0. They were also asked conceptual questions related to content in the textbook and the videos. For example, one conceptual question asked students to "explain" a critical feature of the mean: "Use the fact that the median is the halfway point and the mean is the balance point to explain why an outlier typically affects the mean more than the median." These questions from the textbook were selected to help students both calculate the mean and consider the implications of using the mean (e.g., the mean is a balancing point of a distribution, the residuals from the mean add up to 0, the mean is more affected by outliers).

Here we pause to consider the way that the basic flipped course was designed. The instructor (the third author) had learning outcomes and selected questions from the textbook that seemed to target those particular learning outcomes. Does the basic flipped lesson engage in *practicing connections*? Some of the questions, even in the basic flipped class, could have helped students make connections between concepts (there were plenty of questions that contrasted the mean with the median). But there were many connections that were not made. For example, the questions that asked students to consider the mean as a balancing point and that the residuals sum to 0 were two separate questions. In the PC flipped course, we will return to connecting these implications of the mean to each other. In the basic flipped course, there was some opportunity for struggle, but because these concepts were not scaffolded, most of the students either copied explanations from the textbook or waited until the class discussion to write down their responses. And finally, there was no opportunity for deliberate practice -- to engage with concepts like *balancing* and *residuals from the mean* in more complex and new ways.

Turning to the in-class activities of the PC flipped course, the lesson on the arithmetic mean was also guided by an in-class worksheet. The worksheet was developed by the instructor (found in Appendix C), and included a series of scenario-based questions. Most of the questions were conceptual inquiries, with minimal questions requiring procedural calculations (as these were assumed to be covered by the preparatory out-of-class activities). All of the questions were designed to assist in practicing connections between the central concepts related to the mean and, particularly, focusing on the principal goal of the lesson, which is to understand why the mean is so useful for describing central tendency. To achieve this, questions about the residuals adding up to zero and the mean being the balance point of those residuals were asked in context of one another, and also set in multiple and seemingly dissimilar scenarios (about measuring height by rounding to the nearest foot instead of inch, about cow weight guesses on the internet).

To examine the ways in which the PC flipped lesson was able to encourage practicing connections, we can take a closer look at some of the specific sequences in the lesson. To begin, there are several opportunities to practice making explicit connections between concepts, as the students were guided through varied scenarios that, on the surface, do not appear to be connected, but all have the same underlying concepts embedded within them. For instance, to practice connecting the concept of the mean as the balance point of the residuals, the students begin with a small data set (5, 5, 5, 10, 20) and are given the mean (9) and the median (5). They are then asked to explain how the mean and the median are both trying to describe the "middle" of this distribution in different ways. Furthermore, when explaining how the mean is the "middle" they are asked to consider all the *differences* from the data points to the mean. As they compute all of those differences, it becomes apparent that the negative differences are balanced out by the positive differences and sum to zero, and hence, they begin to connect the idea of the mean as a measure of central tendency with the feature that it uniquely balances the residuals around it. As seen in Figure 1a, we have depicted the data set as a dotplot and shown how the negative residuals balance the positive residuals. Because this is a small concrete data

set, students are able to engage in productive struggle as they grapple with the questions initially on their own and with their peers with minimal guidance from the instructor.
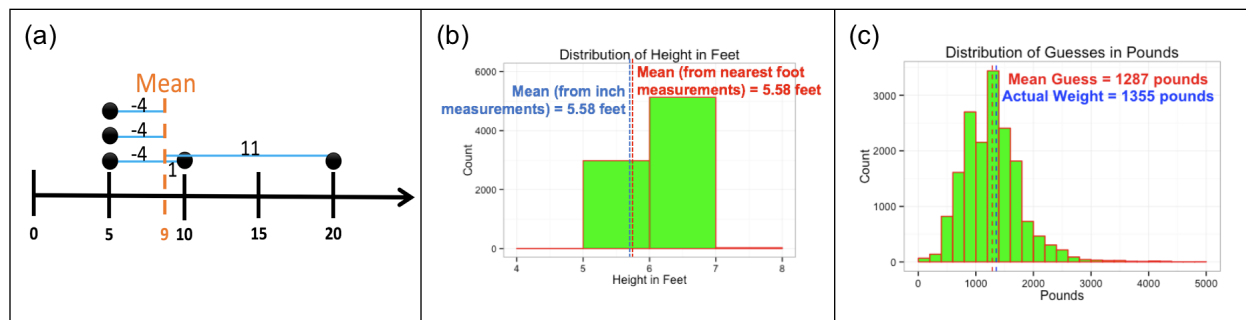

Figure 1. (a) This figure is similar to one that eventually gets developed by students in the course while explaining how the mean is the middle of a distribution.
(b) This figure accompanies the question: "Why is the mean of such awful measurements -- heights rounded to the nearest foot -- so similar to the mean of better measurements -- heights rounded to nearest inch?"
(c) This figure shows the distribution of guesses of how much a cow (in a photograph) weighs. Students are asked to explain why the mean of these guesses are so similar to the actual weight of the cow. All three activities require students to adapt their idea of the mean balancing the residuals to different circumstances in increasingly challenging ways.

This exercise is followed by an example of "measuring heights badly," where students first get practice with trying to visually mark their estimate of the mean on a histogram of heights in inches by considering what point along the x-axis would balance the distribution on either side like a fulcrum. They are then asked whether they would consider measuring height in feet, as in rounding to the nearest foot, to be as good of a measurement as measuring height by rounding to the nearest inch or centimeter. Students will typically respond by saying that rounding to the nearest foot is not as precise of a measurement so it would not be a good way to measure height. The class then tests it out by taking a few students' heights and rounding them to the nearest foot, and then comparing the average of their heights in feet to the average of their heights in inches. Then students are shown an entire distribution made up of these nearest foot measurements and nearest inch measurements. They are shown that the means of the two distributions are quite similar (see Figure 1b). Upon seeing how similar the two means are, and discussing how the times we rounded up were balanced by the times we rounded down, they are encouraged to connect this to the concept of the mean as a balance point of the residuals, as explicitly delineated in the previous example. This provides deliberate practice via an unexpected encounter with this concept, but allows for an explicit connection to be made between the experience and the major learning objective.

Following this class exercise in the worksheet, a new scenario is introduced: a survey that had been conducted by National Public Radio's (NPR) Planet Money. The survey had asked visitors to guess the weight of a cow pictured on their website. The students are also asked to view the picture and take a guess at the weight of the cow. They are also asked to consider things such as: "How accurate could we get by guessing (especially since most people do not know much about cows)? What would you consider accurate?" Students are shown a histogram of the over 17,000 guesses and that the resulting average guess, 1,287 pounds, was pretty close to the actual weight of the cow, 1,355 pounds (see Figure 1c). The students are asked to grapple with the question of why the average guess was so close even though the guesses were individually so terrible. This is now their third attempt to

connect the idea of the mean balancing the residuals to a situation: the overestimations were balanced out by the underestimations. Students are asked to make deep connections between these diverse examples of the mean balancing the residuals: with simple sets of numbers, with terrible measurements, and with guesses from uninformed respondents.

Thus, the goal of the in-class worksheet is to guide students in achieving the advanced learning objectives through practicing conceptual connections across multiple contexts. Yet, pertinently, the connections are not easy for students to make across different contexts. They are presented as seemingly unrelated scenarios with fundamental dissimilarities (e.g., sources of "error" differ), and thus the connections come as a surprise. The practice feels difficulty and thus is an opportunity for deliberate practice. Such practice enhances students' capacity to generalize to new contexts. Embedded throughout the lesson, there are also moments to engage in productive struggle where students struggle answering the questions on their own or with peers, try leveraging what they do understand in a novel situation, and test bad ideas along with good ones. In sum, this lesson incorporates the three features of the Practicing-Connections hypothesis: (1) making explicit connections between concepts, (2) productive struggle in making those connections, and (3) deliberate practice with connections where the tasks are varied and increase in complexity over time. These qualities, combined with a sharp focus on the advanced learning objective, grants students a chance to succeed in achieving that learning outcome; not to repeat rote calculations of the mean or adding up the residuals and getting 0, but to understand more coherently how these concepts connect to a variety of contexts.

Comparing Outcomes from Basic versus Practicing-Connections Approaches

Now we will zoom out from the specific lesson and evaluate these designs at a broader level by comparing the student outcomes for each of these flipped design approaches: basic versus PC. We will compare measures such as exam scores, shifts in mindset about intelligence and math ability, and growth in general quantitative reasoning. The section that follows will attempt to quantitatively assess the efficacy of each of these two approaches so that we can draw more comprehensive conclusions about how to successfully modify a flipped classroom model based on theories of learning.

Method

Course Description and Participants

The course that was flipped was called "*Inferential Statistics for Psychology*" and was considered an upper level statistics course that "provide[d] an introduction to inferential statistics. Topics included: hypothesis testing, probability, parametric (z, t, and F tests, correlation/regression) and nonparametric (Chi-square) procedures, the written presentation of statistical results, and the use of statistical software packages" (from the course catalog). Allotted class time included four lecture hours and three laboratory hours per week. The course is a core course in the Psychology major, meaning that all students in the major are required to take this course. Psychology students commonly cite the difficulty of statistics, lack of use in future career, and dislike of math as contributing to their negative attitudes towards statistics courses (Griffith, Adams, Gu, Hart, & Nichols-Whitehead, 2012).

The third author taught two sections of the course (the basic flipped course in the Fall quarter and the PC flipped course in following Spring quarter of the same academic year). This statistics course is traditionally taught as a lecture and lab course with 24 students; the basic flipped course was taught in this format. The PC flipped course increased enrollment to almost 48 students in a lecture with two

separate lab sections (24 students each). In total, 72 participants were involved in the study. The demographics of the course reflected the base rates of the major and the university.

To enroll in this class, students were required to pass a prerequisite course (descriptive statistics). Combined across the basic and PC flipped courses, 80% of the students reported taking the prerequisite more than one academic year prior to the current class.

Design and Procedures

In this quasi-experimental design, we examined three outcomes between the basic flipped and PC flipped courses. The two types of flipped courses used the same four lab tests and final lecture exam. There were 10 weeks in the term with finals week (we call it Week 11). The lab tests occurred during the lab section at Week 4, 7, 9, and 11. The final exam occurred during Week 11 during the lecture section. Students in both classes were also administered questions about their implicit theories of intelligence and math ability (popularly called mindsets, Dweck, 2007) at both Week 1 and 11. Finally, the PC flipped class was assessed on a measure of quantitative reasoning at Week 1 and 11.

Measures

We used exam scores, a mindset survey, and a test of quantitative reasoning as measures of performance in this study.

Exams. To assess student performance, we examined test grades. There were 5 tests in all: 4 lab tests and 1 essay final exam. The lab tests, administered in a computer lab, were designed to mimic what experts do when confronted with a data set. Students were provided with a data set and a brief description of why and how the data were collected and the variables that were measured. The data and context were all novel situations that had never before been discussed in class, so these were all transfer tests rather than tests of memory. Students were asked 3-5 research questions (e.g., "Does caffeine affect athletic performance?") and were asked to answer them using the provided data set in APA format. These lab tests examined students' ability to connect statistical knowledge to novel contexts. Because they were transfer tests, students were allowed to use any notes or other materials they wanted to bring to the computer lab. This also mimics data analysis in valid situations (most researchers have access to all their notes and the internet when conducting data analysis). Examples of these tests can be found at bit.ly/statsteach. The four lab tests were worth 2%, 4%, 7%, and 14% of the grade for the course. The final exam consisted of 16 short essay questions that asked students to explain various statistical concepts. Example short essay questions are available in Appendix D. The final exam was worth 25% of the course grade.

Mindsets. Students' mindsets about their ability to grow in intelligence and math ability were measured both before and after completing their respective class (the basic flipped or PC flipped course). There were six statements that students rated on a six-point scale (levels of agreement/disagreement). Half of these statements were about intelligence and the other half were about math ability. Intelligence growth mindset was assessed by students' agreement with statements such as "You can always substantially change how intelligent you are" and "You can learn new things, but you can't really change your basic intelligence" (Blackwell, Trzesniewski, & Dweck, 2007). Similarly, math ability growth mindsets were assessed by questions such as "You can change your basic math ability level considerably" and "To be honest, you can't really change how much math ability you have." Negative statements that indicated a fixed mindset were reverse coded such that higher numbers meant more agreement with growth mindset. See Appendix E for all statements (more coding information is provided here: http://bit.ly/mathmindsetquestions).

Quantitative Reasoning. Lastly, quantitative reasoning (based on Stigler et al., 2010) was measured through a 21-question test. This test included almost no calculations, but asked students questions designed to target their understanding of symbols, reasoning with operations, and judgment of magnitude. This test was only administered in the PC flipped classroom design at the beginning and the end of the course. This quantitative reasoning test did not affect students' grades in the course. A copy of the test can be downloaded here: http://bit.ly/quantitativereasoningtest.

Results

Exam Scores

There were four lab tests and a final exam. We conducted ANOVAs to compare test scores between the two flipped classes (basic versus PC) and the results are shown in Figure 2. PC flipped students performed significantly better on the first lab test, $F(1, 70) = 12.01$, $p = .001$, $\eta^2 = .15$ and the final exam, $F(1, 69) = 5.72$, $p = .02$, $\eta^2 = .07$. The two flipped classes were not statistically distinguishable on the second, $F(1, 69) = 3.21$, $p = .08$, third, $F(1,70) = 1.578$, $p = .21$, and fourth lab tests, $F(1, 69) = .11$, $p = .74$.
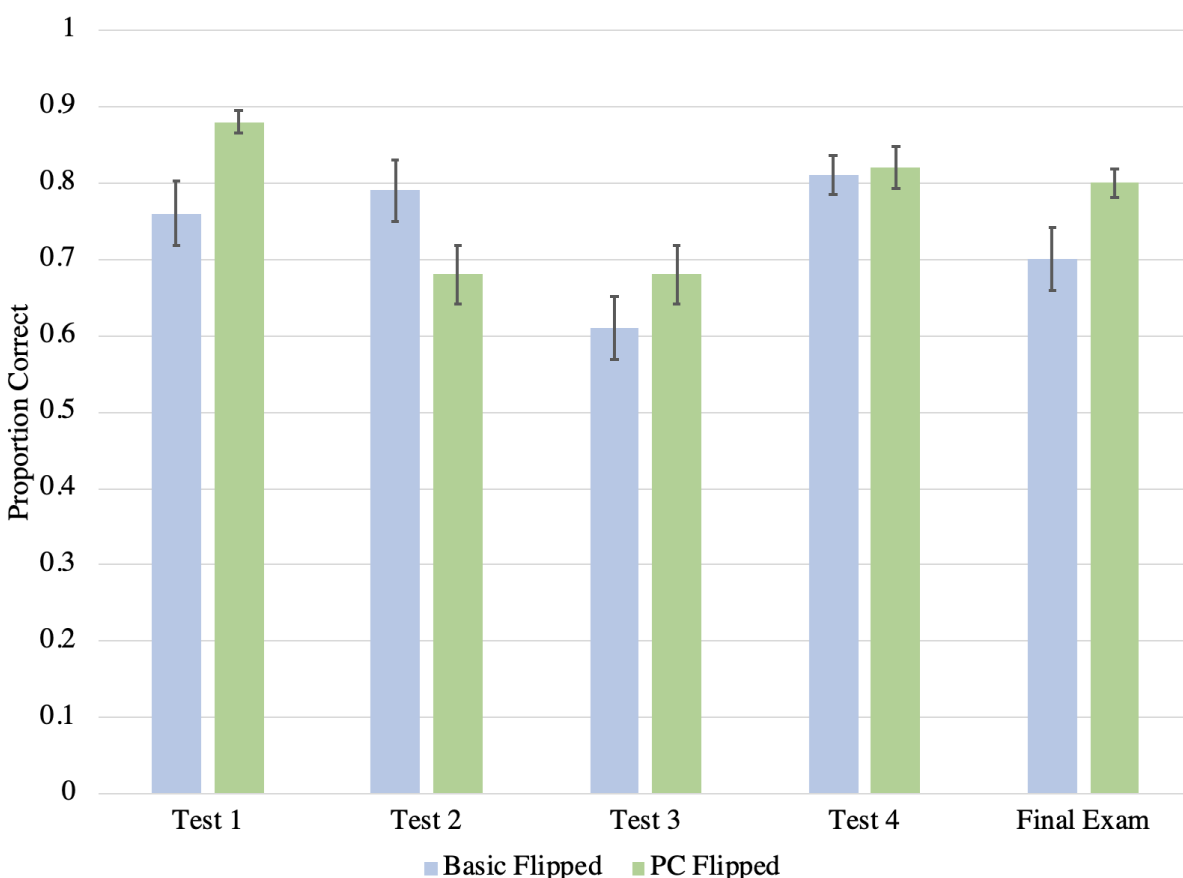

Figure 2. The mean percent correct on each exam for both basic flipped and PC flipped classes.

Mindsets

We calculated mindset change scores by subtracting the pre-class score from the post-class score. A positive score indicates more growth mindset after the course while a negative score indicates a more fixed mindset. A 2 (Flipped Class: basic vs. flipped) x 2 (Domain: intelligence vs. math ability), mixed repeated-measures ANOVA revealed that the type of flipped class had a significant effect on mindset, $F(1,62) = 9.61$, $p = .003$, $\eta^2 = .13$. The PC flipped class experienced a more significant positive shift towards growth mindset in both intelligence and math ability. There was no main effect of domain, $F(1, 62) = 3.35$, $p = .07$, nor was the interaction significant, $F(1,62) = .02$, $p = .89$.
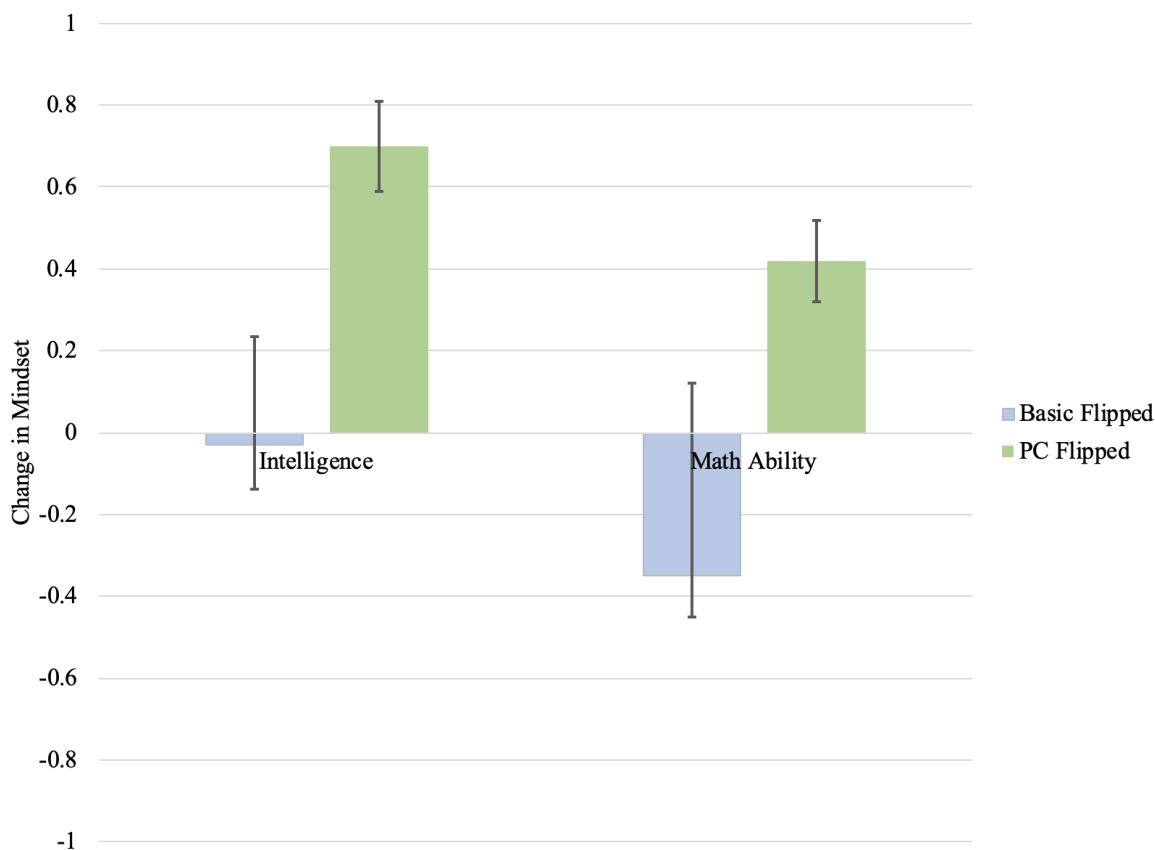

Figure 3. The change in intelligence and math ability mindset for both basic flipped and PC flipped classes.

Quantitative Reasoning

A repeated-subjects ANOVA revealed a significant difference between the beginning of class scores and the end of class scores on the quantitative reasoning exam in the PC flipped class, $F(1, 47) = 12.74$, $p = .001$, $\eta^2 = .21$. This indicates a reliable shift in quantitative reasoning on a test that was not directly related to the course content between the start and finish of the class.
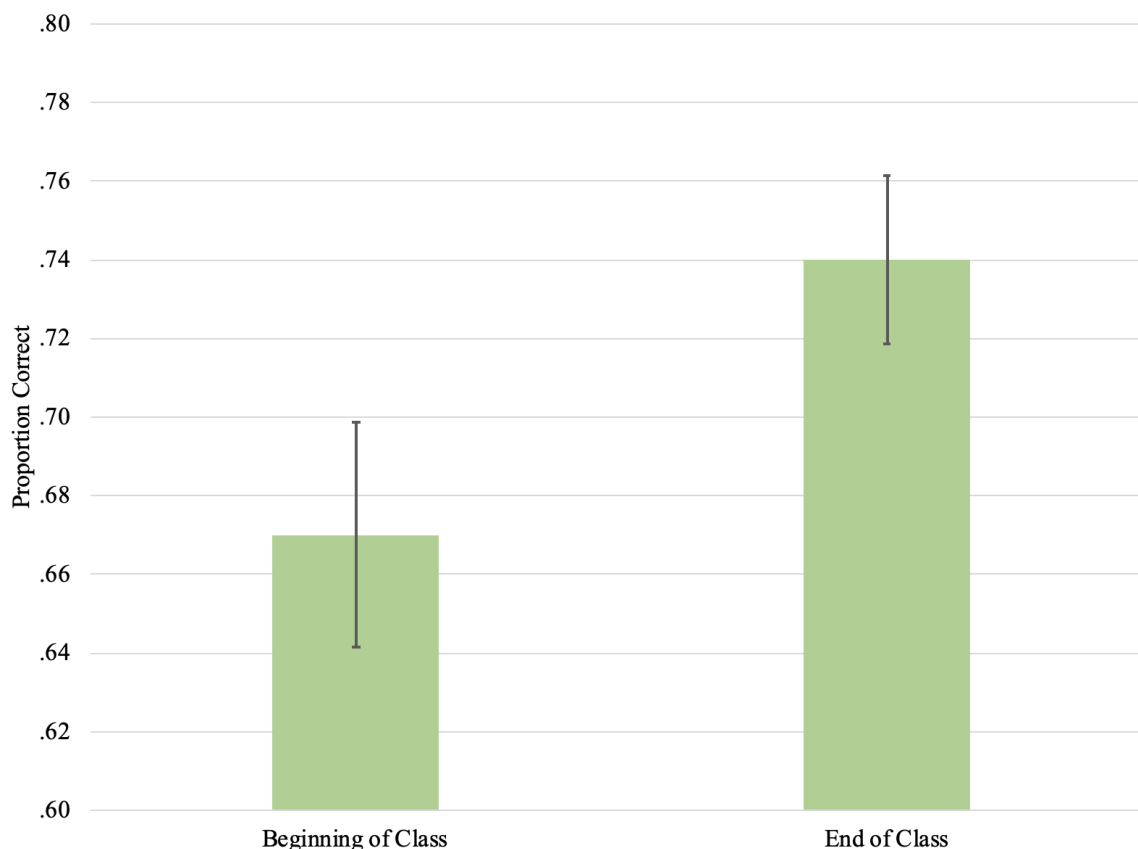
Figure 4. The mean percent correct on quantitative reasoning exam for PC class only.

Discussion

Just as there are better and worse lectures, there are better and worse flipped classrooms. In many ways, efficacy studies compare a nascent flipped classroom to a well-developed culture of lecturing. The time has come to develop a more nuanced view that any pedagogical model is quite broad and provides only the barest outline for possible learner-teacher interaction. There are lots of details that need to be worked out and many of those details may hold the key to better more consistent implementation of flipped models.

In this research, we attempted to develop, implement, and test a set of design principles for flipped pedagogy based on findings in cognitive science. Socioemotional (i.e., mindsets) and cognitive (i.e., course exams and quantitative reasoning) aspects of learning are both important to address in math classes, but mindsets are rarely assessed in a traditional statistics course. The PC flipped class design helped students develop growth mindsets about intelligence and math relative to the basic flipped course. This is an important finding because these mindsets can influence persistence and resiliency in learning and testing (Yeager, & Dweck, 2012). Mindsets may be particularly important in courses that engage students in struggling with difficult concepts.

Although there was a distinct difference in the mindset change between the two flipped classes, the achievement tests only partially demonstrated the benefits of the PC flipped class over the basic one. On a subset of the lab tests and the final exam, which asked students to explain concepts from the course, students in the PC flipped classroom design performed significantly better. Taking a look at the lab tests that did not show a significant difference, we suspect that students needed more practice connecting computer-based activities (using SPSS) with the core concepts of the course. The

PC hypothesis provides a framework to look for missing connections and develop lessons to help students forge those connections in future revisions of course materials. We speculate that the use of software should be interleaved with core concepts earlier in the learning materials. Thus, when students are engaged in lab skills, they will also coordinate them with conceptual knowledge. However, by the final exam, students in the PC flipped course were significantly better at applying concepts from the course because they had practiced making connections rather than actively practicing textbook problems.

Furthermore, students in the PC flipped class increased their quantitative reasoning ability between the beginning of the class and the end of the class. Although this type of reasoning was not addressed directly in class, they were able to show an increased understanding of symbols, reasoning with operations, and judgements on magnitude. Overall, their general ability in mathematics reasoning grew despite not addressing this type of learning directly. However, this result alone is difficult to interpret because the quantitative reasoning test was not administered to the basic flipped class.

What We Learn from Implementing the PC Hypothesis

One of the valuable lessons we can learn from implementing the PC hypothesis is that it provides a framework for an instructional designer's judgment. Whenever an instructor sets out to redesign a whole class (e.g., trying to flip a class), there are innumerable design decisions involved. Because the PC hypothesis focuses on the core instructional goal of engaging students in deliberate practice of making explicit connections, its simplicity, modifiability, and emphasis on deep learning provides a theoretically motivated basis for making many of these decisions. Any usable research-based framework should supply guidance for adapting course materials for different contexts, methods of delivery, and student populations.

The broader goal of using theories of cognitive science to influence course design has been attempted by a number of frameworks. Notable implementation theories include the Knowledge-Learning-Instruction (KLI) (Koedinger, Corbett, & Perfetti, 2012) and Interactive, Constructive, Active, and Passive (ICAP) framework (Chi & Wylie, 2014). These theories offer learning activities best for eliciting a specific cognitive process. Such frameworks offer valuable insight on how to capitalize on the active learning time that a flipped classroom affords and to appropriately align the activities when learning is more passive, particularly, outside of the classroom (since this is an equally important component of a flipped course design). Frameworks like these offer what instructors are often looking for, which is: given some learning outcome, what activities and materials does research suggest will best lead to achieving that outcome?

However, even these highly respected, broadly modifiable, and theoretically motivated frameworks are not often used by instructional designers nor faculty. We speculate that one barrier might be that it takes a significant amount of work to figure out how each learning goal should be achieved in the appropriate way. If these frameworks are used, they are often used with a few lessons or learning outcomes, but not effectively implemented throughout the course. Additionally, since most instructional designers are not cognitive or learning scientists, they may find it more helpful to see an example of a full class implemented with such a framework than to abstractly hear about the framework.

The PC hypothesis takes these considerations and as an accompaniment to this scholarly article, we provide an implementation example in a full set of course materials (available at http://bit.ly/statsteach). These materials can be used in classrooms without the instructor having a background in learning sciences or having a deep understanding of the motivating theoretical framework itself. The theory-driven design makes it possible for materials themselves to embody the

PC hypothesis rather than relying on the hypothesis to be adopted by individual instructors. This makes it possible for instructors to become interested in the learning science-based approach after seeing and using the materials. Even as instructors are gradually learning about the design framework, they can potentially reap the benefits of the embedded emphasis on having students practice making connections.

Although we began with introductory statistics, this framework is general enough for any course. This framework will hopefully engage educators and designers in developing their own judgment for modifying their materials to promote interconnected, flexible knowledge. For instance, a chemistry professor might use the PC hypothesis to consider the key connections that need to be made in their domain. There is considerable research on how connecting multiple representations of nanoscale particles (e.g., connecting particle models to symbols and macroscale phenomena) can help students gain a more coherent understanding of chemistry (Prain, Tytler, & Peterson, 2009; Treagust, Chittleborough, & Mamiala, 2003). The PC hypothesis provides a framework that puts practicing those connections at the focus of instructional design and provides the perspective for varying the connecting activities over time to sustain a level of challenge for the students.

Theories of Implementation for the Flipped Classroom

Recently, there has been an increasing interest in the flipped classroom. However, many of these high quality research programs are difficult to translate into instructional design flexible enough to implement in a different course and in a variety of educational settings because many of them culminate with a test of efficacy ("does the flipped version work significantly better?"). We need more than just tests of efficacy, we need a theoretically-driven design perspective that we can incrementally improve upon. The broader goal of this modest work is to develop a theory of *implementation* so that classrooms can be flipped to maximize student learning.

Our hope as researchers, instructors, and designers of educational materials is to begin with a hypothesis that is able to guide practitioners. If someone reads about a flipped classroom in statistics that was "effective", they would still be left with the question of how to turn the insights from that particular instantiation into concrete actionable designs for their own course content, student population, and institution. We have begun with a few design principles (the PC hypothesis) that we believe can guide the design of course content and instructional routines. Our research goal was to document the instantiation of these design principles to further our understanding of how to flip a course effectively.

Along the way, we have learned that teaching and learning always take place in a system. Flipping the course is only one part of that overall system. Teaching the content of a course cannot be divorced from how we *measure* learning and how we *conceptualize* learning. Our measures of learning must be as current as our teaching techniques. The exams and quantitative reasoning measures used in this study were designed to measure transferable learning and adapting concepts to novel problems and contexts. If we had used traditional exams that ask students about definitions and carrying out procedures (e.g., carry out a t-test), we would not have the data we need to adapt our teaching toward the goal of transferable learning. But we have much further to go in developing assessments that truly measure adaptive understanding of novel situations. Any work in developing better theories and techniques of teaching will be dependent on effective and sensitive assessments.

Another salient lesson is that we need to broaden how we conceptualize learning. As cognitive scientists, we put a heavy emphasis on concepts and mental work. As content experts, we care more about students learning content than about changing attitudes. But all too often, "learning of concepts" is equated to "memorizing" and so practicing concepts is just repeating those concepts

ineffectively. The notion of practice is not applied to attitudes and understanding of concepts. The PC hypothesis attempts to directly address those underlying ideas about learning and practice. The notion of "practice" should be applied to help students make explicit connections in systematically varied and increasingly complex ways. Also, attitudes and socioemotional stances must be practiced. A talk on persevering through difficult problems is less effective than practicing those attitudes in situ. The benefit of the flipped classroom is the apportionment of more class time to practice such perseverance.

Closing the Loop of Theory and Practice

In order to move educational practice forward in a data-driven manner, there needs to be more of a connection between developing learning theories and implementing these abstract ideas in a course. An approach like ours is an attempt to create a bridging type of theory, a set of implementation guidelines that connect basic research in cognition with practice in the classroom. Such an approach can reveal where theories are lacking. For example, although we know that productive struggle routines, where a student attempts a problem before hearing a lecture, can work in a 1-hour experiment (Kapur, 2012, 2014), in a real-life classroom, some students might just learn to wait until the lecture and may cease working on the problems independently once they learn that a lecture will be forthcoming. An implementation theory also reveals how common instructional practice can benefit from basic research. For example, the transfer literature (e.g., Gick & Holyoak, 1980, 1983; Gentner, Loewenstein, & Thompson, 2003) has an extensive set of findings that can help us develop assessments that can measure flexible learning.

There are several other approaches to create bridges between basic research and instruction that have influenced this work (Koedinger, 2002; Merrill & Twitchell, 1994). For instructional practice to move forward, we need studies like this one that attempt to design a whole course, from start to finish, based on these principles. Depending on the field and content area, different approaches might provide different benefits. Instructors and practitioners need to see more examples of these theories implemented in classrooms. Ultimately, theory and practice should work in a loop where theories help guide instructional design, and implementation can illuminate issues in need of better theories.

Conclusion

By incorporating the PC hypothesis into a flipped classroom, we found that there were improvements in mindsets, course learning, and quantitative thinking. These results are promising, but there is also room for improvement. Our framework illuminated areas in our course that were still lacking connections for students (e.g., a disconnect between computer-based SPSS experience and core statistical concepts). Not only does this framework help identify areas that may be lacking, but it also guides us in our understanding of what we could do better in future classes (e.g., make more explicit connections between computer-based experiences and core statistical concepts). In our research, we attempted to improve upon the basic flipped course through a design framework grounded in modern theories of learning. Merely implementing a flipped method classroom will not automatically affect student learning. There is nothing magical about the format of flipping itself. It is the process of incrementally improving the flipped classroom that can incrementally improve learning.

As John Dewey said in *Sources of Science Education* (1929), "No conclusion of scientific research can be converted into an immediate rule of educational art. The value of the science… resides in the enlightenment and guidance it supplies to observation and judgment of actual situations as they arise (p. 19, 32)." Much of instructional design must be guided by the expertise of educators about their

context, their content domain, and their students. The PC hypothesis allows that expertise to work together with what we know from research on teaching and learning to result in guiding design decisions as they arise.

Acknowledgements

Appendix

Appendix A. Quiz (corresponding with assigned videos) for Basic Flipped Class

1.   For this set of numbers [1,1,1,6,6] what is the:

   Mode? _____

   Median?   _____

   Mean?   _____

2.   What does this symbol (Σ) indicate?

3.   What is the symbol for the mean of the sample?

4.   What is the symbol for the mean of the population?

5.   There were 11 people in a room who made an average salary of 25K per year. Another person walked in the room. That person was unemployed and made $0 per year. Which of the following statements below is true?
   a.   The mean for all 12 people is smaller than the mean of 11 people.
   b.   The mean for all 12 people is larger than the mean of 11 people because the sum of 12 numbers is bigger than the sum of 11 numbers.
   c.   The mean for all 12 people is larger than the mean of 11 people because 12 is bigger than 11.
   d.   None of the above can be verified without knowing what each person's salary is.

6.   In this set of numbers, the mean is equal to the median [1,1,3,5,5]. Change one number in the set so that the mean is greater than the median.

   _____ _____ _____ _____ _____

7.   Range and inter-quartile range are usually used with which of the following measures of central tendency?
   a.   Mean

    b. Median
    c. Mode
    d. None of the above

8.   Standard deviation and variance are usually used with which of the following measures of central tendency?
    a. Mean
    b. Median
    c. Mode
    d. None of the above

9.   Which of the following statements are true?
    a. The sum of all squared deviations from the mean = 0.
    b. The absolute value of a deviation from the mean = 0.
    c. The sum of all deviations from the mean = 0.
    d. All of the above are true.

10.   What is the symbol for sample standard deviation?

11.   What is the symbol for population standard deviation?

12.   What is the "sum of squares"?
    a. Sum of squared values
    b. Sum of squared standard deviations
    c. Sum of squared distances from the mean
    d. Sum of means
    e. Sum of squared means

13.   What is the conceptual difference between s and S?
    a. s is the estimate of the sample standard deviation from the population distribution and S is the sample standard deviation.
    b. s is the estimate of the population standard deviation from the sample distribution and S is the sample standard deviation.
    c. s is the sample standard deviation and S is the population standard deviation.
    d. s is the population variance and S is the sample variance.

14.   Which of the following statements is generally true?
    a. $S = s$
    b. $S > s$
    c. $S < s$
    d. $S = \sigma$

Appendix B. Basic Flipped Lesson on Arithmetic Mean

1. Find the mean and median of each ordered list, and contrast their behavior. (Feel free to use excel!)
    a. 1, 2, 3   Mean _____   Median _____
    b. 1, 2, 6   Mean _____   Median _____
    c. 1, 2, 9   Mean _____   Median _____
    d. 1, 2, 297   Mean _____   Median _____

2. Use the fact that the median is the halfway point and the mean is the balance point to explain why an outlier typically affects the mean more than the median.

3. Five children (about 4 ft tall) are standing together when their teacher (who is 6 ft tall) joins the group. What is the new mean height? The new median height?

4. Verify that the sum of the deviations from the mean is 0 for the numbers 1, 2, 4, 6, 9. Find the standard deviation. Explain what that number means.

5. Without computing, match each list of numbers in the left column with its standard deviation in the right column. Check any answers you aren't sure of by computing [you can use excel functions if you wish].
    a. 1 1 1 1   i. 0
    b. 1 2 3   ii. 1.581
    c. 1 2 2   iii. 1
    d. 10 20 20   iv. 0.577
    e. 1 2 3 4 5   v. 5.774

6. The mean of a set of seven values is 25. Six of the values are 24, 47, 34, 10, 22, and 28. What is the 7th value?

7. The sum of a set of values is 84, and the mean is 6. How many values are there?

Essay Question Reflection (Connect to Explanations in Lecture Videos)
8. In inferential statistics, we usually are talking about s (rather than S or σ) when we refer to standard deviation. Why is this the case?What is the conceptual difference between s and σ? (The formulas are not necessary for this answer.)
9. How are the concepts of mean and standard deviation similar? How are they different?
10. How does the formula for s relate to the definition of standard deviation?
11. In the formula for s, why do we divide by the term n-1 rather than n?

Lab practice: Introduction to SPSS
We will be looking at a data set introduced on pg. 33 (Display 2.17): Life expectancies in countries around the world. Use SPSS to explore the shape, center, and spread of this distribution.

1. How to convert excel into spss data file
    a. Download **lifeexpectancies.xlsx** (either from "Notes" in course website or from google drive folder) and convert to an .sav file.

b. What is the difference between an .sav file and an .spv file?
c. In general, which one do you need to turn in?
2. How to create a quick histogram in SPSS (use "Graphs → Legacy dialogs → Histograms")
    a. Why does this histogram look different than the one on pg. 33 (Example 2.3)?
3. How to get common descriptive statistics in SPSS (there are a few different ways but let's explore "Analyze → Descriptive statistics → Frequencies")
    a. Mean, median, mode
    b. Standard deviation and variance
    c. Minimum and maximum
    d. Range and interquartile range (well… at least the quartile cutoffs)

Appendix C. PC Flipped Lesson on Arithmetic Mean

Summary Statistics vs. Data Visualizations

Data often comprises many, many numbers. Imagine a study with 10,000 participants. It's hard to keep in mind the results of each participant. Summary measures (such as the mean) are easier to keep in mind. Data visualizations (such as histograms) are easier to explore and think about.

Sometimes people get hung up on the summary measures; these summaries can hide telling patterns. Sometimes people think: all we have to know is the mean and that will be enough information about the *distribution*.

Let's say you are trying to choose between two cancer drugs. Which of these is a better drug?
Here are two cancer drugs. Which is a better drug? (Assume the number means additional months lived.)

Drug A: 23, 21, 5, 9, 18, 24, 6, 8, 8, 25, 23, 12 (mean is 15.17)
Drug B: 17, 16, 16, 13, 15, 22, 17, 13, 15, 15, 9, 15 (mean is 15.25)

If we just looked at the means, we would think that drug B might be slightly better (the difference expressed as .08 of a month turns out to be about 2.5 days).
1. Why is drug A better than drug B?
2. Why is drug B better than drug A?
3. Imagine someone looked at the means and summarized the two drugs like this: "Drug B basically helps you live 2.5 days longer than drug A." Is this an accurate statement? Why not?

Histograms vs. Bar charts

Figure 1.	Figure 2.
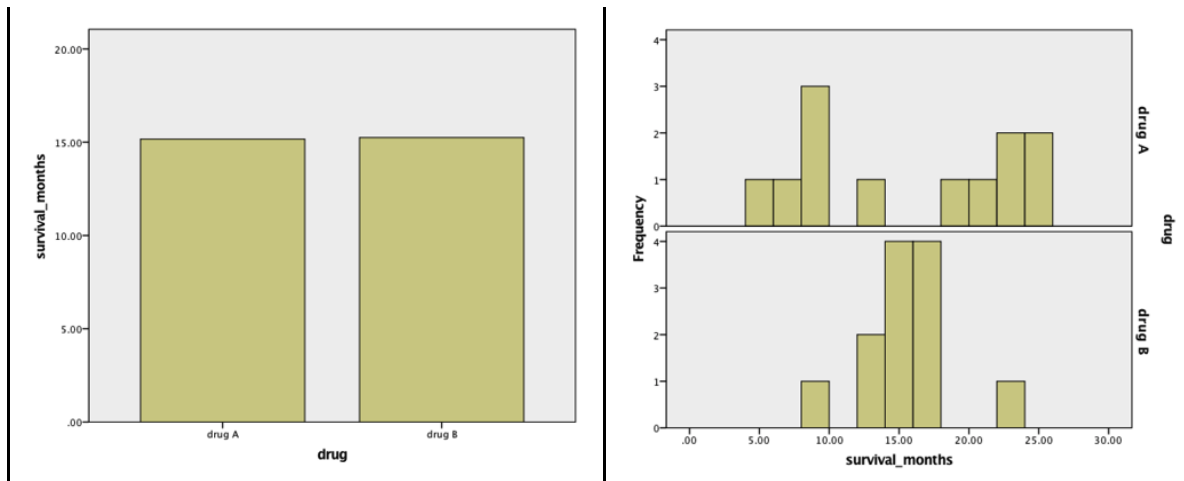
4. Which figure depicts the bar chart? Which is the histogram? How did you know?
5. Which one would be more helpful to a person who is trying to figure out which drug to take? Why?
6. Now remember that the mean for drug A is 15.17 and 15.25 for drug B. On each of the figures above, put a mark where the means should go. Where do the means go on a bar chart? Where does it go on the histogram? How are they different?
7. Let's say the pharmaceutical company tells people, "Well, on average, people live about 15 more months with this drug."
    a. Is that an accurate statement for each drug?
    b. Which drug is better described by the mean? Why?

Figure 3.	Figure 4.
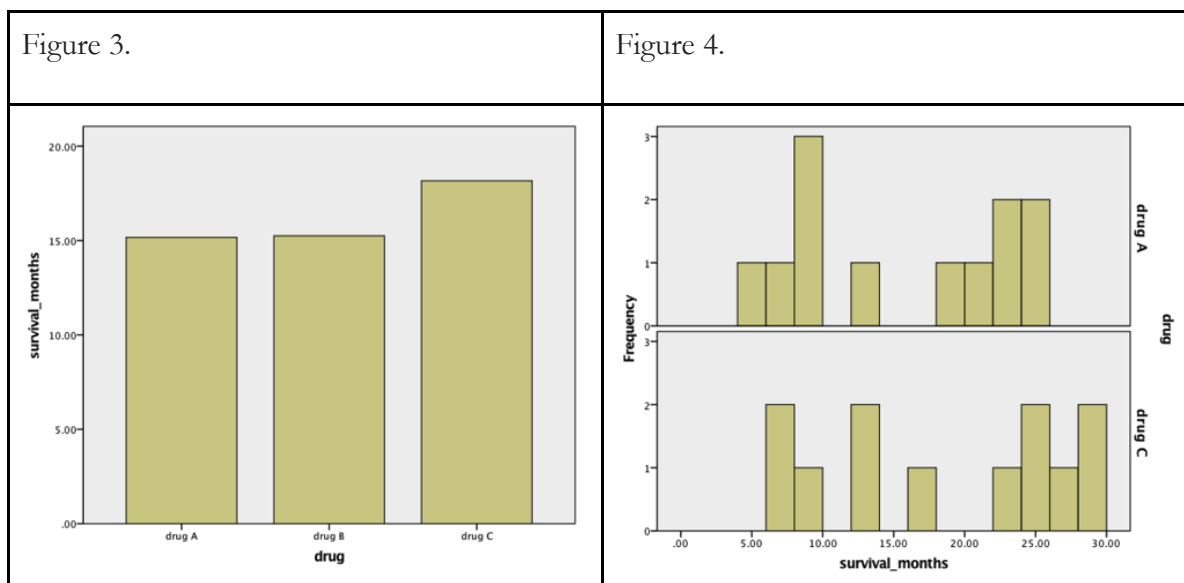
8. A new drug comes out on the market. Let's call it drug C. Take a look at the bar charts and histograms above. How is it similar to drug A? How is it different from drug A?
9. Without knowing anything else, which seems like the better drug to take (assuming you want to survive a longer period of time)? Why?
10. When it is enough to look at the bar chart that summarizes the distributions? When is it more

helpful to look at distributions with histograms?

Pros and Cons of Means

11. You have probably heard that medians are better description of central tendency than means for some types of data. When/why are medians better than means?
12. Even so, we are going to be using means a lot in inferential statistics! You should wonder – why? If medians are sometimes better than means, why are means used so much? What is a mean good for? We are going to cover this in the rest of the lesson today.

What does middle mean?

13. Here is a small distribution: 5, 5, 5, 10, 20. The median of this distribution is 5 but the mean of this distribution is 9. Both are trying to describe the "middle" or "central tendency" of this distribution.
    a. How is the median (5) the "middle" of this distribution?
    b. How is the mean (9) the "middle" of this distribution? (Hint: Consider all the *differences* from the data points to the mean…)

Measuring Heights Badly

14. This is some data from the National Longitudinal Study of Youth. They collected data from 8000+ youths (15-20 years old). Here is a histogram of their heights. Mark on the histogram about where you think the mean might be.

Distribution of Height in Inches
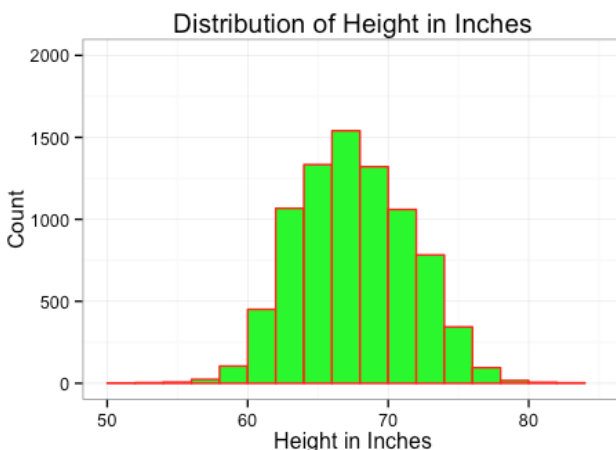
15. The average is 67.05 in (or 5.59 ft). Was your mark near that value? Adjust if necessary.
16. What if we only had measures to the nearest foot? Are these measurements just as good as measures to inches? Why or why not?
17. If we took the *average* of measures to the nearest foot, how accurate would the average be? (To be clear, that's like if we took someone who is 5' 3" and rounded them to 5' and took someone 5' 11" and rounded them to 6' and then took the average of the 5' and 6'… the average of these two people would be 5.5 ft.)
18. Here is the distribution of heights measured in nearest foot. Do you think the average is similar or different from the average height measured in inches?
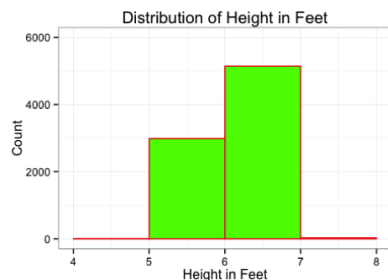
19. The mean (based on the rounded heights) is _____. Why do you think the two means are _____, even though the measurement of each individual's height is so much less exact?

People who know nothing about cows

Photograph Credit: David Kestenbaum/NPR

Let's take a new situation. NPR (Planet Money) posted this picture of Penelope the cow, and asked people to guess how much the cow weighed. (For scale, the guy weighs 165 lbs.)

14. How much do you think the cow weighs? Write down your guess.
15. How far off do you think your guess is? Do you think it's too high? Too low? Not sure?
16. How accurate could we get by guessing? What would you consider accurate?
17. Here are the 17000 guesses collected on the internet. Why do we see all this variation? (This next figures were revealed to the class on the projector.)
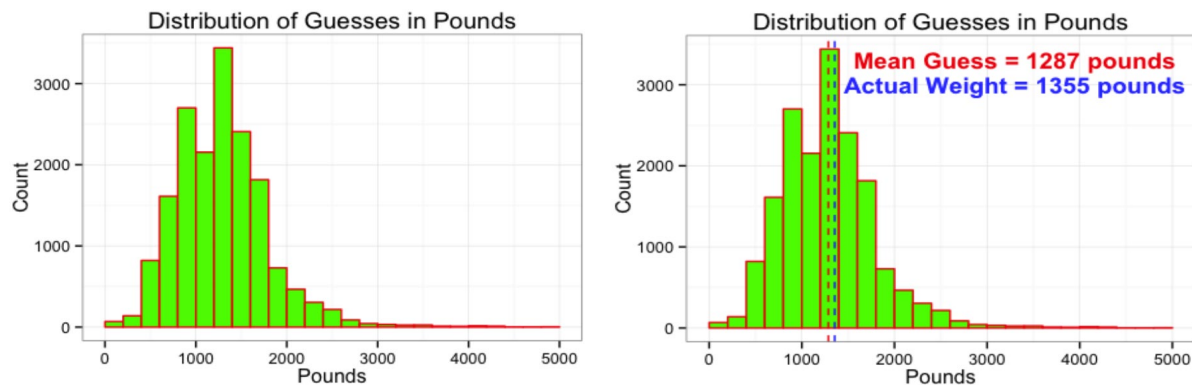
18. Here's the crazy part. It turns out the average of all these guesses, is pretty close to the actual weight of the cow! Why might this be surprising (whether you personally are surprised or not)? Why is the average so close to the actual weight of the cow?
19. Why are the measurements "bad" in this situation?
20. Why is the average "good" in this situation?
21. What is the advantage of the mean (over the median) based on our discussion today?

Appendix D. Example Final Exam Questions

How are the concepts of mean and standard deviation similar? How are they different?

What is the conceptual difference between s and σ? (The formulas are not necessary for this answer.)

Why is the sampling distribution important to inferential statistics?

What is the Central Limit Theorem (CLT) and what is its purpose?

What are the similarities and differences between hypothesis testing and confidence intervals?

Appendix E. Mindset Questions

The statements that are italicized were reverse coded because high responses indicated a fixed (or entitative) mindset regarding intelligence/math ability. Strongly disagree was coded as a 1 and Strongly agree was coded as a 6. Students' responses for intelligence and math ability were separately averaged. High scores indicate more growth (or incremental) mindset.

Intelligence Statements
You can always substantially change how intelligent you are.
You can learn new things, but you can't really change your basic intelligence.
You have a certain amount of intelligence, and you can't really do much to change it.

Math Ability Statements
To be honest, you can't really change how much math ability you have.
You can change even your basic math ability level considerably.
No matter how much math ability you have, you can always change it quite a bit.

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
100

References

Austin, J. L., Lee, M. G., Thibeault, M. D., Carr, J., E., & Bailey, J. S. (2002). Effects of guided notes on university students' responding and recall of information. *Journal of Behavioral Education, 11*(4), 243-254. https://doi.org/10.1023/A:1021110922552

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers and Education, 78*, 227-236. https://doi.org/10.1016/j.compedu.2014.06.006

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Science, 22*(4), 637-660. https://doi.org/10.1017/S0140525X99532147

Barsalou, L. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes, 18*(5-6), 513-562. https://doi.org/10.1080/01690960344000026

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological science, 16*(2), 101-105. https://doi.org/10.1111/j.0956-7976.2005.00789.x

Bergmann, J. & Sams, A. (2012). *Flip your classroom: How to reach every student in every class every day.* Washington, DC: International Society for Technology in Education.

Blackwell, L. S., Trzesniewksi, K. H. & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and intervention. *Child Development, 78*(1), 246-263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Burke, A. S. & Fedorek, B. (2017). Does "flipping" promote engagement?: A comparison of a traditional, online, and flipped class. *Active Learning in Higher Education, 18*(1), 11-24. https://doi.org/10.1177/1469787417693487

Chase, W. G., & Ericsson, K. A., (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition.* Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in Chess. *Cognitive Psychology, 4*, 55-81. https://doi.org/10.1016/0010-0285(73)90004-2

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219-243. https://doi.org/10.1080/00461520.2014.965823

Cohen, P. A., Ebeling, B. J., & Kulik, J. A. (1981). A meta-analysis of outcome studies of visual-based instruction. *Educational Technology Research and Development, 29*(1), 26-36.

Damasio, A. R. (1989). 3. Concepts in the brain. *Mind & Language, 4*(1), 24-28. https://doi.org/10.1111/j.1468-0017.1989.tb00236.x

Deslauriers, L., Schelew, E., Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031). https://doi.org/10.1126/science.1201783

Dewey, J. (1929). *The sources of a science of education.* New York: Liveright.

Dweck, C. S. (2007). *Mindset: The new psychology of success.* New York, NY: Ballantine Books.

Ericsson, K. A., Krampe, R. T., Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406. https://doi.org/10.1037/0033-295x.100.3.363

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Galway, L. P., Corbett, K. K., Takaro, T. K., Tairyan, K., & Frank, E. (2014). A novel integration of online and flipped classroom instructional models in public health higher education. *BMC*

Medical Education 14, 181. https://doi.org/10.1186/1472-6920-14-181

Garofalo, J. (1989). Beliefs and their influence on mathematical performance. *The Mathematics Teacher, 82*(7), 502-505.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*(2), 393-408. https://doi.org/10.1037/0022-0663.95.2.393

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38. https://doi.org/10.1016/0010-0285(83)90002-6

Gilboy, M. B., Heinerichs, S., Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior, 47*(1), 109-114. https://doi.org/10.1016/j.jneb.2014.08.008

Glenberg, A. M. (1997). Mental models, space, and embodied cognition. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes* (pp. 495-522). Washington, DC, US: American Psychological Association.

Griffith, J. D., Adams, L. T., Gu, L. L., Hart, C. L., & Nichols-Whitehead, P. (2012). Students' attitudes toward statistics across the disciplines: A mixed-methods approach. *Statistics Education Research Journal, 11*(2), 45-56.

He, Y., Swenson, S., & Lents, N. H. (2012). Online video tutorials increase learning of difficult concepts in an undergraduate analytical chemistry course. *Journal of Chemical Education, 89*, 1128-1132. https://doi.org/10.1021/ed200685p

Hill, J. L., & Nelson, A. (2011). New technology, new pedagogy? Employing video podcasts in learning and teaching about exotic ecosystems. *Environmental Education Research, 17*, 393-408. https://doi.org/10.1080/13504622.2010.545873

Jensen, J. L., Kummer, T. A., Godoy, P. D. d. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE-Life Sciences Education, 14*, 1-12. https://doi.org/10.1187/cbe.14-08-0129

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379-424. https://doi.org/10.1080/07370000802212669

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science, 40*, 651-672. https://doi.org/10.1007/s11251-012-9209-6

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science: A Multidisciplinary Journal, 38*, 1008-1022. https://doi.org/10.1111/cogs.12107

Koedinger, K. R. (2002). Toward evidence for instructional design principle: Examples from Cognitive Tutor Math 6. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G., Wiegel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of Twenty-Fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Vol. 1* (21-49). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive science, 36*(5), 757-798. https://doi.org/10.1111/j.1551-6709.2012.01245.x

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1), 30-43. https://doi.org/10.2307/1183338

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
102

achievement in mathematics. *Journal for Research in Mathematics Education, 30*(5), 520-540. https://doi.org/10.2307/749772

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science, 341*(6149), 976-980. https://doi.org/10.1126/science.1238041

Mayer, R. E. (2001). *Multimedia learning.* Cambridge, UK: Cambridge University Press.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. *US Department of Education.*

Merkt, M., Weigand, S., Heier, A., Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction, 21*(6), 687-704. https://doi.org/10.1016/j.learninstruc.2011.03.004

Merrill, M. D. & Twitchell, D. G. (1994). *Instructional Design Theory.* Englewood Cliffs, NJ: Educational Publications, Inc.

Neef, N. A., McCord, B. E., & Ferreri, S. J. (2006). Effects of guided notes versus completed notes during lectures on college students' quiz performance. *Journal of Applied Behavior Analysis, 39*(1), 123-130. https://doi.org/10.1901/jaba.2006.94-04

Nouri, J. (2016). The flipped classroom: For active, effective and increase learning - especially for low achievers. *International Journal of Educational Technology in Higher Education, 13*(33), 1-10. https://doi.org/10.1186/s41239-016-0032-z

Prain, V., Tytler, R., & Peterson, S. (2009). Multiple representations in learning about evaporation. *International Journal of Science Education, 31*(6), 787-808. https://doi.org/10.1080/09500690701824249

Schoenfeld, A. H. (1989). Explorations of students' mathematical belief and behavior. *Journal for Research in Mathematics Education,* 338-355. https://doi.org/10.2307/749440

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14*, 293-305. http://dx.doi.org/10.1016/j.learninstruc.2004.06.005

Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator, 1*(3), 4-16.

Stigler, J. W. & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York, NY: Free Press.

Son, J. Y., Ramos, P., DeWolf, M., Loftus, W., & Stigler, J. W. (2018). Exploring the practicing-connections hypothesis: Using gesture to support coordination of ideas in understanding a complex statistical concept. *Cognitive Research: Principles and Implications, 3*(1), 1-13. https://doi.org/10.1186/s41235-017-0085-0

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4

Treagust, D., Chittleborough, G., Mamiala, T. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education, 25*(11), 1353-1368. https://doi.org/10.1080/0950069032000070306

VanLehn, K. (1989). Rule acquisition events in the discovery of problem solving strategies. *Cognitive Science, 15*(1), 1-48. https://doi.org/10.21236/ada225579

Watkins, A. E., Scheaffer, R. L., & Cobb, G. W. (2011). *Statistics: From data to decision* (2nd ed.) John Wiley & Sons, Inc.

Wilson, S. G. (2013). The flipped class: A method of address the challenges of an undergraduate statistics course. *Society for the Teaching of Psychology, 40*(3), 193-199. https://doi.org/10.1177/0098628313487461

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
103

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*(4), 302-314. https://doi.org/10.1080/00461520.2012.722805

Yong, D. Levy, R., Lape, N. (2015). Why no difference? A controlled flipped classroom study for an introductory differential equations course. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 25*, 907-921. https://doi.org/10.1080/10511970.2015.1031307

Fame and Fortune: Developing a Simulation Game for the Music Industry Classroom

Monika Herzig
Indiana University
mherzig@indiana.edu

Abstract: This study documents the development of a classroom game simulating the effects of contractual arrangements on the economic relationships between artists, record labels, and consumers. The game was tested with multiple revisions in classroom settings over a period of three years by using surveys, interviews, session videos, and teacher observations. Using the approach of Grounded Theory, observations and insights were extracted from the collected data with the goal of identifying effective strategies for developing and using classroom simulations. Findings include evidence of deeper engagement with the subject through cognitive, psychomotor, and affective learning.

Keywords: classroom game, simulation, engagement, motivation, music industry

Introduction

Rieber and Noah explain that "Games are a way of knowing the world, a mediation between experience and understanding" (2008, p. 79). Malaby (2007) offers a compelling definition of games as "a semibounded and socially legitimate domain of contrived contingency that generates interpretable outcomes." Hence, games are about disorder (Huizinga,1971; Caillois, 2001) providing a set of rules that guide the process of engaging in a game towards a set of predictable and unpredictable outcomes. The opportunity to mimic everyday experience during the gaming process allows for interpretation of the various outcomes as contingencies of factors encountered during the process. A discussion of what happened and what could have been otherwise contingent on the randomness of the game components, the social unpredictability of fellow players, actions performed with various grades of success by the players, and even the unpredictability of meaning allow for compelling life-like learning in the
classroom (Malaby, 2007, 1999).

The study of games is often linked to play as a subcategory and thus creates a separation from the 'real' world. Generally, the following attributes are associated with play (a) it is usually voluntary and enjoyable; (b) it is intrinsically motivating, that is, it is pleasurable for its own sake and is not dependent on external rewards; (c) it involves some level of active, often physical, engagement; (d) it is distinct from other behavior by having a make-believe quality; (e) the outcome is uncertain; (f) it is unproductive in that the activity does not produce goods of external value; and (g) it is governed by rules (Blanchard & Cheska, 1985; Caillois,1961; Csikszentmihalyi, 1990; Pellegrini, 1995; Pellegrini & Smith, 1993; Yawkey & Pellegrini, 1984). But as Taylor (2006) documents, the separation of play from real life when studying games can't be upheld and creates a marginalization of the scholarship of gaming.

Crookall, Oxford, and Saunders (1987) distinguish between games and simulations, where simulations attempt a representation of some real-world system in contrast to games being their own "real" systems. Nevertheless, the cost of error or losing is low in both, simulations and games, except in games such as poker where real financial assets are at stake (Garris et al, 2002). Roleplay is a

subcategory of the simulation games. The goal is to act out 'real-world' behavior while keeping the impact of errors on the 'real world' of other people low. Within the defined parameters of a role in a given scenario, the participant is allowed to make strategic choices thus allowing some exploration of alternative strategies in acquiring the desired skills from the given scenario.

There is little consensus though in the literature on the definition of a game beyond description of characteristics. Juul (2005) defined games as having six features: (a) a rule-based formal system, (b) variable and quantifiable outcomes, (c) different assigned values for different outcomes, (d) an outcome influenced by the efforts that the player exerts, (e) players feeling emotionally attached to the outcome, and (f) consequences of the activity that are negotiable.

Research on the effectiveness of using games in the classroom offers a wide spectrum of results concerning learning outcomes, motivation, and skill acquisition. While motivation and engagement in the activity itself is usually high, few games have empirically demonstrated that they can successfully teach academic content when used in the classroom on a standalone basis (Garris et al., 2002). Qualitative literature reviews also revealed the importance of pedagogical guidance by the teacher during classroom games. Kangas, Koskinen, and Krokfors (2017) identified the teacher's role as pedagogically active in the majority of studies reviewed, specifically for planning game-based, learning, and integrating game-based learning into teaching after the gaming situation. The teacher should define what content knowledge students are expected to learn by engaging in the game and make links within the curriculum (Barab et al,
 2010). Throughout the game process, teachers should guide learning through specific questions focusing on the concepts (Arnab et al., 2013; Barab et al., 2010; Chee & Tan, 2012; Watson et al., 2011).

In order to determine learning outcomes in game design, Bloom's Taxonomy of Teaching and Learning (1956) is a useful measurement tool. Research has provided some evidence for the successful acquisition of complex tasks through the use of educational games (Garris & Ahlers, 2001; Ricci, Salas, & Cannon-Bowers, 1996). Serious games have been proven to increase motivation and the student's desire to learn (Leach, G. J., & Sugarman, T. S., 2005, McClarty et al, 2012) with curiosity and the desire to complete goals being major factors.

The Indiana University Center for Innovative Teaching and Learning (CITL) sponsored a faculty group in 2012/13 exploring the concepts of intrinsic motivation and play in the classroom. Every member in the group developed a game to be used in their classroom teaching and the games were tested in classroom settings as well as with colleagues. Using the results of a variety of evaluation tools such as surveys, video analysis, and interviews, the games were revised for optimal learning and motivational experiences. Based on the research discussed in the faculty learning group and using my expertise in music, I developed a simulation game for the music industry. Following is a report on the development, implementation and optimization phase, as well as learning outcomes.

The research model employed for this study is based on the principles of
 ethnomethodology and grounded theory, an approach suggested by Matthew Sharritt (2011) to improve game content and interface and achieve meaningful and serious outcomes. Grounded theory, originally developed by Barney Glaser and Anselm Strauss (1967), is a method of comparative analysis with the goal of discovering a theory from data rather than confirming preconceived expectations. Similarly, ethnomethodology allows ideas and principles to emerge through analysis of situation-specific behavior and responses, thus providing the framework for grounded theory.

Fame and Fortune – A Simulation Game

My initial motivation for researching classroom games was the search for an option to simulate the economic and social relationships between artists, record labels, and consumers in a classroom setting.

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
106

The music business has unique challenges and product characteristics that are difficult to recognize in general business simulations and case studies. Such characteristics include the unpredictability of emotional responses, fashions, publicity and other factors related to the artist, product, economy, culture, and society that influence the success of an artistic product and career. In addition, the industry has experienced fundamental restructuring over the past decade due to new technologies such as file sharing and digital recording and transmission tools as well as new ways of communication due to social networking and subscription services, among many others.

As an active musician myself, I have certainly experienced the ups and downs of the music industry and the amount of work and investment that it takes to create a musical product and create a fan community. For my students though, these concepts are more difficult to grasp as their experiences in the industry are mainly as consumers and casual participants. Our society is focused on a superstar artist image resulting in a "Winner Takes it All" economic model. For example, about 1% of touring artists receive close to 50% of all concert ticket revenue in the popular music market, and 80% go to the top 5% of touring artists (Connolly & Krueger, 2006). This lottery-type environment of marketing an artistic product as well as the daily grind involved in being an artist may be easier to grasp with the help of an interactive tool rather than textbooks. Furthermore, chances for success could increase exponentially with the opportunity of testing a variety of real world decisions without real world consequences as often as desired, similar to flight simulators, in order to optimize goal-setting and decision-making.

In order to provide such real-life experiences, many music industry programs around the country have created student record labels that function as a laboratory for aspiring music entrepreneurs. Very few though have been successful in creating meaningful product and sustaining long-term engagement and growth of the student-run business. Common hurdles are legal issues within a university environment, young and inexperienced employees, ownership of intellectual property issues, constant personnel turnover, time limitations due to semester length, and similar issues. Thus, the goal of providing realistic experiences can rarely be achieved with such student-run businesses. Of course, completing internships in a variety of music corporations are also effective ways of learning. But again, such experiences vary in depth and scope depending on internship placements, availability, and geographic location as well as effective mentorship by business leaders who may have very little pedagogical training.

Hence, the option of creating an experiential simulation game seemed ideal. Gredler (1996) described an experiential simulation as creating a particular psychological reality with participants taking on defined roles within that reality. Thus learners interact with a simulation in order to understand the underlying model without having to reach a certain goal state (de Jong et al., 1999). In this case, students take on the roles of artists, record label executives, or consumers and engage in the process of selling a musical product, concert events, and merchandise with the goal of experiencing the factors that influence financial as well as artistic success. Such simulations have been proven successful as players bring their own already-learned rules into an activity. Thus the simulator becomes more of a trigger activated by the actions of the participants within a framework of rules. Cookall et al (1987) argue that simulation situations actually may become more real than ordinary situations as participants enter with a level of uncertainty, testing out actions and their consequences without fear of failure.

As discussed in the introduction, engaging in a gaming situation is a process that creates a set of predictable and unpredictable contingencies capable of generating a variety of meanings (Malaby, 2007). Thus the learning outcomes for this simulation game include the predictable components of learning factors that influence the success of a musical career as well as experiencing the effects of luck, social and economic tendencies, and other random events that shape meaning in unpredictable ways and require problem solving skills and adaption. In order to assess these learning outcomes

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
107

students completed reflection activities prior and after playing the game analyzing the factors that influence an artist career as well as shared their interpretations of meaning during the debriefing exercises immediately following the game.

Game Mechanics

Similar to other economic reality simulation games such as Monopoly or The Settlers of Catan, I chose a combination of action cards, resource cards, value cards in form of simulated money, and interactive rounds with chance factors and final value calculations to determine the outcomes and levels of success. The main learning goals were to demonstrate the effect of specific contract stipulations between artists and their team members, specifically a record label, the behavior of consumers, market constraints of limited budgets, and chance factors influencing market behavior such as prominent performances, viral success, human failure, or new technologies.

Common game features were established during the initial design process.

Hardware

Several sets of cards were designed to establish the roles and rules for the game:

Role Cards - for the initial setup, students draw cards determining their roles as artists, record labels, or consumers. The ideal ratio is five consumers per pair of artist and A&R representation of a label (short for Artist and Repertoire and usually in charge of finding artists and developing their products)
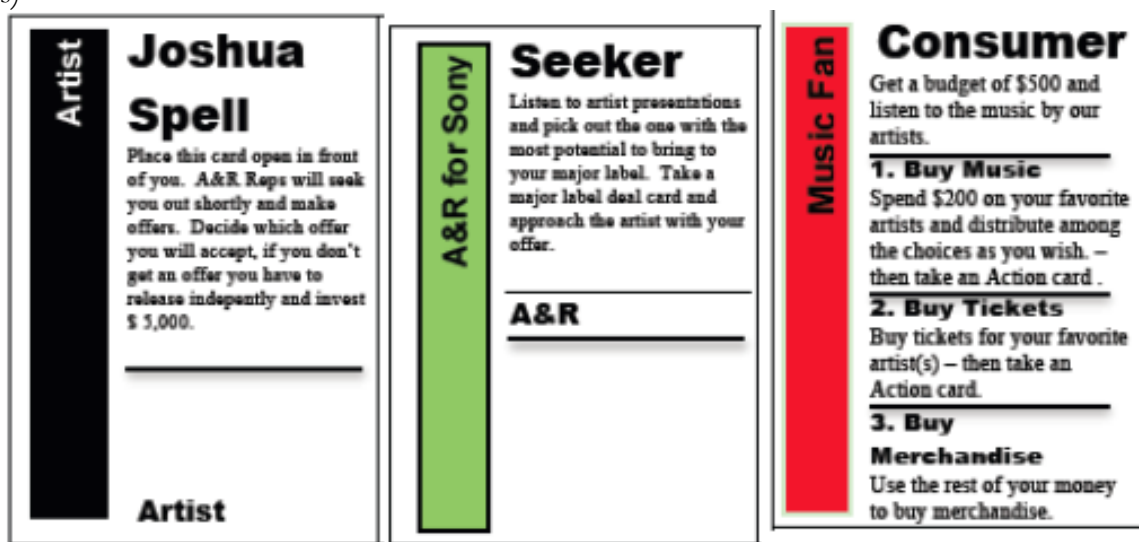
Contract Clauses – the players with A&R roles draw clause cards depending on their status of representing a major or independent label that define the contract options that they can offer to artists

Clause

Major Label Deal Card

Artist will receive a $ 5,000 recording advance and 20% artist royalties from product sales. Artist will pay label 30% of ticket and merchandise sales and keep the rest.

The $ 5,000 advance is a loan (negative income for the artist) and will be recouped from product sales only.

Clause

Independent Label Deal

Label pays production costs of $ 3,000 and radio campaign for $ 1,000. Artist will receive 50% after production costs only are recouped. The label will recoup the radio costs from product sales - all ticket and merchandise sales go to the artist.

Consumer Action Cards – consumers interact with the music industry in ways beyond spending money on product by engaging with the artists on various social levels, these cards reflect these different levels of engagement

Blog

Consumer Action

Your blog about one of the artists goes viral.

Consequence

Bring this news to one of the artists of your choice and the artist adds an extra $ 2,000 in music sales to be distributed with the label according to contract. You'll get some of the artist's merchandise in return for a chance to win.

Manager

Consumer Action

You decide to enter the music business and become the manager of your favorite artist.

Consequence

Team up with your favorite artist and join the team. You will help promote the artist and receive 20% of artist's revenue and now have the opportunity to win the game if your team succeeds.

Artist Action Cards – similarly, artists encounter various opportunities and barriers beyond the basic process of practicing their craft and cards were designed to reflect these factors.

Ticket Cards - the most important income source for contemporary musicians are ticket and merchandise sales at live events, hence cards simulating tickets at various price points were designed reflecting typical price levels for Do-It-Yourself, independent, and major label artists.

Money Cards – play money is exchanged to simulate the financial transactions during the game. Artist/ Label teams use a board and markers to document their financial transactions and income progress. Various candy or other food items serve as merchandise to be sold during the last game round.

Software or Rules

- The rules of the game are set up at the beginning:
- Roles are initially drawn by chance from the role cards
- Each artist has to establish an artist persona at the beginning of the game

- A&R representatives have to find an artist that agrees to their contract terms or become a consumer
- Each consumer receives a fixed budget and has to spend a specific amount of their budget during each round of playing – first round on product, second round on concert tickets, third round on merchandise
- Between each round of play consumers and artists have to draw an action card and follow instructions on the card
- The artist/ label teams have to account their income according to the contract terms defined on the clause cards after each round of playing.
- Any artist may choose not to partner with a label and engage in the game by themselves by initially going $5,000 in debt and keeping all income. However, they will only have access to the lowest tier of concert tickets as well as lower income levels from the chance cards.

Strategy and Evolution

Within the rules, players are allowed varying degrees of choice. Initially, artists choose their persona according to their own preferences and have the opportunity to brand themselves by showcasing a video and advertising their character. The label representatives choose the artist they want to collaborate with based on their own perception of potential. Consumers may spend their total money budgeted for each round on as many artists as they like with any amount they prefer. The chance cards provide options for applications, i.e. consumers may win extra money in the lottery and can distribute the extra budget according to their preferences. Teams may engage in creative advertising strategies throughout the game.

Termination, Goal, and Winning

The goal for the artist/ label teams is the highest financial gains. The goal for the consumers is to influence the financial outcomes according to their musical preferences. After all consumers spent their budget during the third round, the game is terminated and the artist/ label teams engage in their final accounting to determine the winning artist and the winning label.

Competition/Cooperation

The artist/ label teams or do-it-yourself artists compete for the attention of the consumers and their willingness to spend their budgets on their products, concert tickets, and merchandise. Their ability to lure consumers to support their team directly influences financial gains. Some consumer chance cards actually provide the option of joining an artist team, thus increasing the collaboration options for higher impact.

Chance

While the contract terms, label affiliation, and the ability to attract costumers are the highest predictor of success, the two chance cards may derail an artist career on occasion through personal drama,

substance abuse, downturn of the economy, injury, or provide additional financial resources such as high-profile performance, sponsors, or film deals.

Game Setup

The session starts with an initial discussion on what students perceive as factors influencing an artist career based on an assignment submission and a review of common record contract terms. In order to understand the accounting process, the teacher also guides the class through an example of calculating income streams for all parties involved. Then roles are randomly assigned to the participants by drawing cards. After initially introducing their product through Youtube videos and an elevator pitch describing their music and brand, artists are approached by record label executives with cards outlining common contract details for major and independent label deals. Artists agree to get 'signed' according to contract details offered by the A&R representatives or to pursue a career without a label by starting with a $ 5,000 debt that needs to be recouped.

Consumers then engage in three stages of spending their allotted budgets of $500 with $200 on recordings, then $200 on concert tickets, and finally $100 on merchandise (in form of candy or other edibles) from the artists that they hold concert tickets for. Between each purchase round there are random occurrences drawn via cards that influence income, roles, and future endeavors of consumers as well as artists, such as highly visible engagements, social networking fads, loss of band members, etc. Artist/ Label pairs also need to calculate their respective income according to contract terms after each round and display their financial progress.

In order to differentiate between budget sizes and distribution networks that major labels versus independent labels versus independent artists can offer, the initial contract clauses, chance cards, and ticket sales differentiate amounts according to label affiliation, i.e. ticket prices and potential sales figures. The artist/ label partners conduct final calculations after the merchandise purchasing round in order to determine final financial status for each party. All participants then proceed to discuss their observations on what factors shaped the outcome and their learning experiences.

The level of reality of a simulation game is determined by the way participants reduce the level of uncertainty about themselves and the situation by helping shape the constructed reality with their own social experiences and background. Hence, the debriefing process is crucial to allow participants to reflect and examine on what happened and to export the learning and insights gained into their real world (Cookall et al, 1987). At the end of the game, students reflect first on what happened, why some moved ahead financially while others got left behind and the effect of the contract terms. Then the discussion extends towards the various factors that influence financial success in the music industry, reflections on the process of making decisions on how to spend money and who to support, and general emotional and rational revelations of being in the various roles. The initial list of influential factors on an artist career is now expanded with the additional factors discovered as a result of the gaming process. Finally, the lesson concludes with thoughts on how each party can influence their success and navigate chance factors as well as thoughts on implementation of the game experience into real life.

Implementation Process

Over the course of three school years, the game was played for one 75-minute session each semester as part of the Music Industry I curriculum and various related courses in the Arts Administration program at Indiana University with class sizes ranging from 15 – 50 students for a total of 15 times. Feedback was collected from voluntary class surveys and personal interviews, videotaping of the

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
112

sessions, as well as instructor observations. Various revisions were implemented and evaluated throughout the process based on the feedback:

Consumer Engagement

Initially the role of consumers was to spend their budgets as they pleased for each round of spending hence influencing the economic status of each artist. Especially in larger classes, the students in the consumer roles quickly expressed little engagement with the game due to their limited involvement. While artists and labels were busy calculating their income and promoting their products, consumers became passive observers. In order to achieve equal active engagements, consumer chance cards were introduced after each round that facilitated additional actions and even optional change of roles. Also, consumers were assigned to observe one artist/label team in a fishbowl-type setting keeping track of progress by taking notes.

Contract Refinements and Calculations

Over the past decade, the artist-record label relationship has gone through drastic changes due to digital distribution, streaming, new recording techniques, and increased focus on live performance. Traditionally, contracts included negotiated percentage terms from recorded product sales only. Contemporary agreements also ask for percentages and support structures for touring income and merchandise sales. The results are all-inclusive '360' contracts that are much more complex in terms of income and rights distribution. Similarly, the contract terms of the game cards were adjusted to reflect the variety of options for major, independent, and do-ityourself artists. The calculation of income is complex and the process was refined using large displays with involvement by the whole class for most effective learning. Initially, at least one calculation example needs to be completed with the whole class before engaging into the game, here is an example:

     Contract Term Card: Artist will receive a $5,000 recording advance and 20% artist royalties from product sales. Artist will pay label 30% of ticket and merchandise sales and keep the rest. The $5,000 advance is a loan (negative income for the artist) and will be recouped from product sales only. In the calculation below, the team received $4,000 in product sales during the first round. 20% of those sales equaling $800 will go towards the artist debt, meaning the artist now owes $4,200 to the label. In the meantime the $4,000 sales income go to the label leaving only $1,000 unrecouped from the initial investment. During the second round the team receives $4,000 from ticket sales. According to the contract, the artist gives 30% to the label, equaling $1,200 and keeps $2,800 in earnings. The label has now recouped its initial investment and made $200. For the final round, the team sells merchandise worth $1,000. $300 is the label share and

the artist keeps $700. For the final accounting, the artist made $3,500 from tickets and merchandise but still owes the label $4,200 on future product sales to be recouped from the 20% artist share. The label made their investment back and is $500 in the black. Any future sales and percentages from tickets and merchandise are earnings.

Artist	Label	Total
Recouped $800, still $4,200 in debt	- $1,000	Product sales $ 4,000

$ 2,800	$ 1,200	Tickets $ 4,000
$ 700	$ 300	Merchandise $ 1000
- $ 4,200 on product sales, but made $ 3,500 from tickets and merchandise	$500	

To clarify the process, a few additional scenarios should be discussed that may be encountered through chance cards:

What if

> *a. The concert tour stalls in the middle – subtract half of concert ticket income, no merchandise sales.*

> *b. The artist adds an extra $ 1,000 in music sales to be distributed with the label according to contract.*

The table above should be drawn on large boards or projected from computers for each team with total income and calculations filled in after each completed round. Ideally consumers assigned to artist-label groups will assist in calculations and check on accuracy.

Introduction of Artists

The main products in this game are the artists and their music. Initially the artist needs to attract labels and choose most favorable contract terms and subsequently attract consumers to spend their music budgets on their products and concert tickets. Hence establishing an attractive brand and active promotions throughout the game are crucial. Initially, artists were asked to take on a fictive persona and present a short musical demo. This turned out to be extremely intimidating especially for players without musical background. Performing a song of choice with a Karaoke track was also difficult due to time constraints and varying musical backgrounds. A time-effective and realistic solution is having the artists take on a musical persona of their choice and introduce them to the class via a short video demo and an elevator pitch explaining their work, what they have accomplished so far, and why they are attractive brands.

Budgets

Similar to Monopoly game mechanics, the initial budget was set to $500 as starting capital to be spent over the three rounds. It quickly became clear that a certain budget amount/ player ratio is needed in order to produce realistic financial outcomes. In addition to observing an approximate 5:1 ratio for artists to consumers, budgets should be increased for smaller groups.

Class Discussion

In terms of student engagement, session videos document the high level of activity during the game as artists and labels realize that they can move ahead of the income curve by advertising their products and as consumers trade in their money and action cards. But as Cookall et al (1987) documented, the immediate debriefing clarifies and strengthens the learning outcomes. Initially open-ended discussions often with limited time were shaped towards an ideal time period of 15 minutes with the following guidance questions:

> *How did the winners get ahead?*
> *How realistic are the contract terms?*
> *What factors really make the difference to fame and fortune?*
> *What worked well with the game?*
> *Improvement ideas?*

Additional written reflection assignments ensured equal engagement in the debriefing process. Throughout the game, consumers create a written log of their observations of an artistlabel team and all participants complete a short reflection on factors influencing an artist career as observed during the simulation (see assignment instructions and results below).

Results

Cognitive Learning

In order to better understand students' knowledge and perception of the factors shaping an artist career prior and after playing the game, students were asked to submit written reflections at the beginning of the semester and after playing the game. Initially, each student submitted a list of factors in response to the following question:

> *Create a list of any possible factors that you can think of that influence the financial and popular success of an artist career, i.e. talent, label support, effective management, etc. Provide a short argument/ description for each factor on why and how you think this is influential and put them in a possible rank order. Please be as comprehensive as possible with a minimum of five factors but most likely a larger number. We will revisit the list at the end of the semester and reevaluate after learning more information about the industry.*

After the game session the student responded with a written report to the following questions:

How did playing Fame and Fortune and what you learned about recording contracts change your perception of the factors that influence an artist career and/or possibly make you aware of factors that you hadn't considered previously? List new factors that you would add to your initial list.

> *How will playing Fame and Fortune and learning about recording contracts affect your music consumption habits and active engagement with the music industry?*
> *How does playing Fame and Fortune and learning about recording contracts affect the way you'll approach your work in this course?*

Results provide a glimpse on how engagement in the simulation process helped students learn the range of influential factors on an artist career based on the social, economic, and psychological contingencies in the game and real life beyond their initial perceptions from classroom learning and readings. The summary below documents the responses of sixty students who completed both essays

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
115

during the Spring 2014 semester. They initially named a total of 30 different factors that can be grouped into two broad categories: Attributes and skills of the artist (the vast majority) and support resources available to the artist. Most notable is the initial focus on any factors that can be controlled by the artist.

Table 1. Perceived factors influencing the financial and popular success of an artist career prior to playing the simulation game

ARTIST SKILLS/ ATTRIBUTES	RESOURCES
Talent and Performance Skills (69)	Strong Support Team (39)
Perseverance, Drive (47)	Promotion/ Marketing (29)
Uniqueness, aesthetic, appearance (42)	Access to Funds, label support (28)
Passion, genuine, honest, confidence (38)	Luck/ Timing (9)
Network, fan relations (31)	Touring (4)
Management skills, professionalism, Strategic planning (26)	Producers, Engineers (3)
Artist image (19)	Radio Play (3)
Diversity/ versatility/ adaption (11)	Family and Friends Support (3)
Other: Communication skills, legal knowledge, attractive themes, tech knowledge, consistency, sellout, sacrifice, forgiveness, compassion, awards	Location (2)

After playing *Fame and Fortune* the same sixty students completed their reflection essays and acknowledged the following factors as new additions to their initial lists as a result of playing the simulation game.

Table 2. Additional factors influencing the financial and popular success of an artist career after playing the simulation game

ARTIST SKILLS/ ATTRIBUTES	RESOURCES
Strength/ weaknesses of personal life (19)	Effect of contract terms (32)
Artist Branding (15)	Luck/ Market (26)
Unique Product (8)	Record label business models (9)
Commitment/ Perseverance (4)	Consumer support in Superstar society (9)
	Proprietary rights/ control (8)
	Access to funds (7)

The comparison documents the shift in student perception from a linear relationship between artist skills and career success to realizing the multiplicity of interrelationships between many stakeholders and the influence of economical, social, and cultural factors. In addition, 68% of the students (n=41) indicated that the simulation helped them to better understand the reality of artist-label relationships and 75% (n=45) expressed the realization of their influence as consumers and indicated new consumption habits to better support the artists of their choice. Seventeen students pledged to buy more merchandise, seven acknowledged that they will change their consumption habits, four decided to pay for music, others pledged to be more active and support more independent labels and artists, and found new appreciation for the work of an artist, while eleven claimed no change in their consumption habits. This realization of the consumer influence is remarkable as none of the students had indicated this specific factor in their initial lists.

The final debriefing sessions after playing the simulation documented the cognitive learning outcomes for the students. Similar to the written reflections, students expressed their revelations about the effects of contract terms, the financial realities of the industryy, the importance of fan support, the influence of social media and consumer behavior, superstar effects, having to deal with chance factors and other outside influences beyond individual control.

Furthermore the videotaped classroom sessions document the heightened engagement during the interactive segments through increased volume level, chatter, laughter, gestures and movement. Once the teams realize that they can capture buyers by advertising their goods, they strategically explore the parameters of the simulated world as described by Crookall et al (1987).

In contrast to lecture sessions, students rarely engaged with their electronic devices during the game sessions and commented on how they enjoyed getting to know their classmates and being able to focus better by moving around in interviews and surveys. Students do find it difficult to concentrate over extended periods of times during lectures and frequently experience anonymity resulting in decreased engagement in large-lecture settings (Blatchford, Edmonds, & Marin, 2003; Fenollar, Roman, & Cuesta, 2007, Young, Robinson, & Alberts, 2009).

Logs from two different game sessions were coded for Teacher Talk, Discussions, Student-led activities, and Movement Activities. The time spent moving while interacting with artists and purchasing goods was logged with 13.5 minutes and 12.5 minutes respectively which constitute 26% and 21% of the total instruction time. Furthermore, the amount of teacher talk was limited to 17 minutes (32.7%) and 21.5 minutes (35.8%) mostly providing game instructions, while discussions took up 13.5 minutes (26%) and 12.5 minutes (20.8%) and student-led activities, mostly introducing themselves as artists and sharing their economic results took up 7 minutes (13.5%) and 13.5 minutes (22.5%). It should be noted that the second session was a larger class with five rather than three artist/label teams hence the introduction of artists took up more time.

Table 3. Classroom Activity Time (in minutes)

	Session 1	Session 2
Total Time	52	60
Teacher Talk	17	21.5
Discussion	14.5	12.5
Movement Interaction	13.5	12.5
Student Leadership	7	13.5

On the other hand, the social activity and movement can create a chaotic atmosphere and for some students the movement and chatter can be intimidating. Especially in larger classes some students requested clear examples to avoid confusion during the interactive segments and focused class attention during the chance card segments. The quick pace and limited time for instructions also received criticism for not allowing ample time to understand the process and feeling disorganized.

Affective Learning

The initial goal of developing classroom games in the faculty group was to study the effect of play on intrinsic motivation and engagement with the subject. Ruohomäki (1995) claimed that simulation games provide attitude changes and increased motivation and interest towards the subject matter. Sixteen participants during the Fall 2014 game session completed a voluntary survey after the game session and responded all positive to the question "Overall I found this game enjoyable".

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
117

Herzig

Strongly Agree	4 respondents	24 %	
Agree	12 respondents	71 %	
Undecided		0 %	
Disagree		0 %	
Strongly Disagree		0 %	

When asked if the game made them more motivated to learn 78 % of the students also responded positive.

Strongly Agree	1 respondents	6 %	
Agree	12 respondents	71 %	
Undecided	2 respondents	12 %	
Disagree		0 %	
Strongly Disagree		0 %	
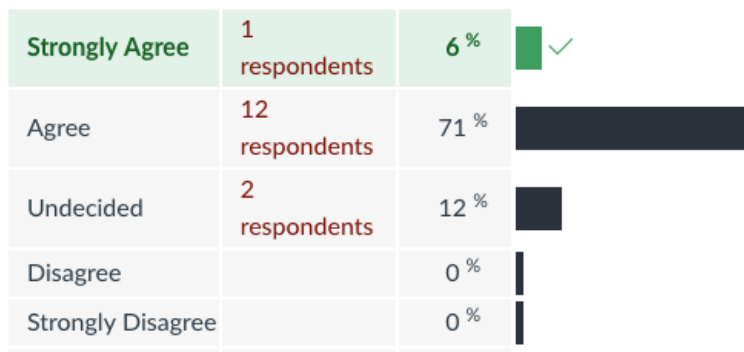
53% of the students indicated that the game helped them remember course materials better.

Strongly Agree	1 respondents	6 %	
Agree	8 respondents	47 %	
Undecided	6 respondents	35 %	
Disagree		0 %	
Strongly Disagree		0 %	
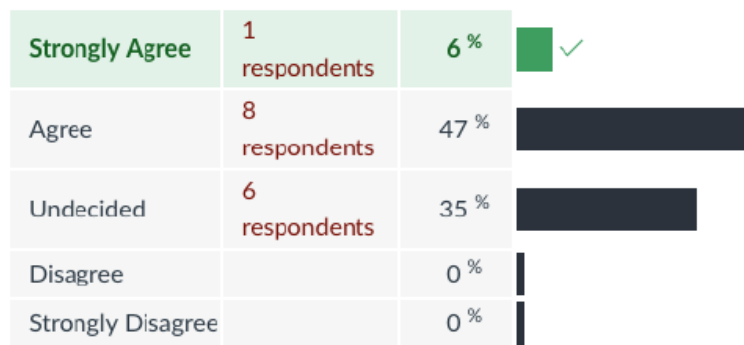
Responses to the question "I'd like to use games for other concepts in this class" were also 83% positive.

Strongly Agree	3 respondents	18 %	
Agree	11 respondents	65 %	
Undecided	1 respondents	6 %	
Disagree		0 %	
Strongly Disagree		0 %	
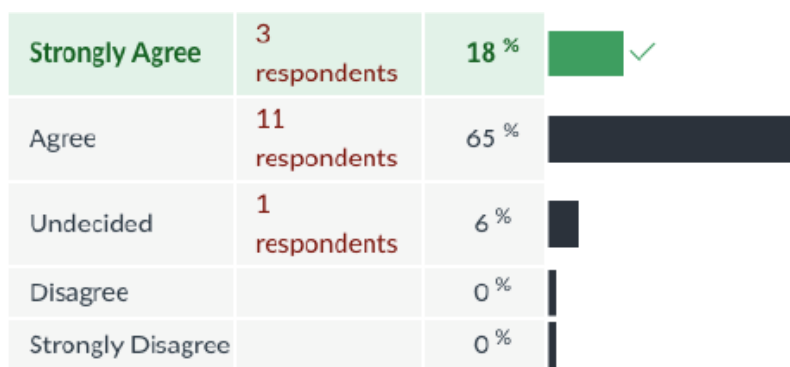
As documented earlier, more than 80% of the students who completed the written post reflection indicated changing their behavior and attitudes towards the industry, mainly by becoming paying and more conscious consumers. Similarly in the interviews, students indicated increased enthusiasm and curiosity towards the subject matter. Comments mentioned especially the effect of the active learning on their engagement during the session and subsequently with the subject matter in contrast to having to listen to lectures.

Figure 2. Engagement Plot during Video Session 1
Level 0 = class introduction, Level 5 = full student attention, Level 10 = lively interaction
Figure 3. Engagement Plot during Video Session 2
Level 0 = class introduction, Level 5 = full student attention, Level 10 = lively interaction
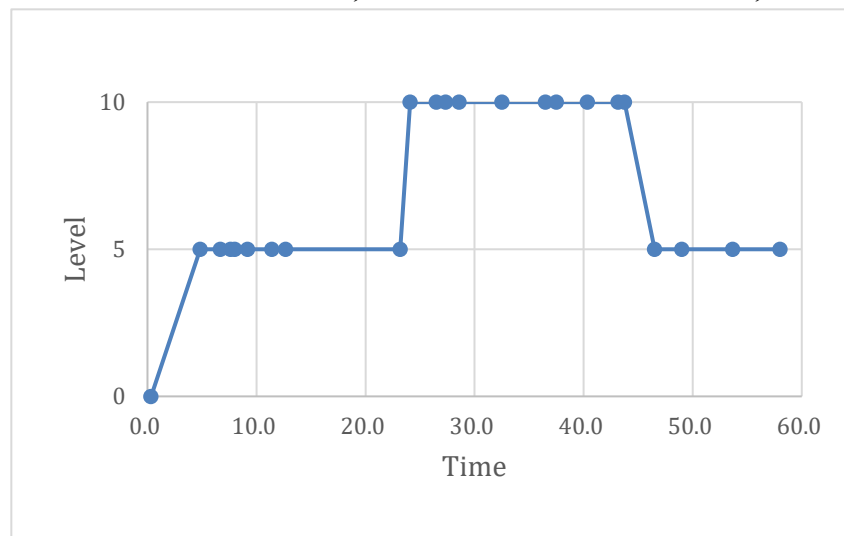
Studies found a strong relationship between students' ranking of their performance in the game and their reported enjoyment and perception of acquired knowledge and generalizability of the game (Remus, 1977; Remus and Jenner, 1981). Similarly, student interview and survey responses indicated stronger engagements for those who took on artist roles versus students who played as consumers in a more passive way. For example, one of the students who impersonated Beyoncé commented: "I thought it was fun especially since I won it so I was like, this is cool, I should do it again". On the other hand, a student in consumer role mentioned: "And then the part I didn't really like was I was a consumer so I felt like I wasn't as involved in the game as if you were an artist or a record label because they were the ones that were really affected by the cards and what happened in the game. Where if you're a consumer you're just kind of going along with it all. You didn't really have a choice".

Overall students affirmed increased engagement with the subject matter in their survey responses after the playing session. For example, 15 respondents from the Fall 2014 class session responded with 71% agree and 6% strongly agree with two undecided on the question if the game made them more motivated to learn. They all agreed (with two strongly agree) that playing the game made the course content more fun.

Discussion

This study on designing and developing a classroom game offers various insights and opportunities. Overall, responses were positive in terms of increased engagement and motivation as well as depth of

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
119

learning and practical applications. Using the process of ethnomethodology and grounded theory, an approach suggested by Matthew Sharritt (2011) the results and observations form the basis for the following recommendations.

The design of the game requires detailed analysis of the desired learning outcomes, the influencing factors and parties involved, as well as studying various game mechanics in order to determine the ideal approach. Experienced game developers highly recommend an initial paper and board prototype before considering electronic options. The analysis presented in the study by Wilson et al (2009) on the relationship of game attributes and learning outcomes provides useful guidelines for optimizing learning outcomes according to Bloom's taxonomy (1956). The desired outcomes will also determine the ideal match of using a simulation versus a game (Gredler, 1996).

Effective game development requires the development of a prototype through multiple stages of trials, refinements, and effective evaluation as described for example in Taspinar, Schmidt, and Schuhbauer (2016). In this example, refinements in game mechanics, content, as well as process are documented during the three-year implementation process and evaluation and feedback during each session provides further refinement options. Ideally, implementation of classroom games is a dynamic process with a continuous evaluation process in place.

This open mode of classroom interaction is unusual and creates discomfort for some students in comparison to the traditional teacher-guided classroom lectures and discussions. Thus it is important to start the session with detailed explanations of the goals and process in order to facilitate positive engagement and a structured environment. Also an instructor-guided final debriefing is crucial for effective learning as documented by Cookall et al (1987) and confirmed throughout this study. Follow-up activities will strengthen the increased engagement with the subject and may consist of further examples, readings, applications, and reflections. In this case, the collection of additional factors learned by playing the game through the debriefing as well as reflection exercises demonstrated the expanded scope of learning.

Overall, active teacher involvement is an essential component of student engagement and learning when using classroom games (Kangas et al, 2017; Arnab et al., 2013; Barab et al., 2010; Chee & Tan, 2012; Watson et al., 2011). In this study, uncertainty and confusion mostly resulted from a perceived lack of teacher guidance and/or instructions. The valuable feedback is helpful for further revisions of the game mechanics in this case and should be considered when designing a classroom game.

The use of a simulation game as an opportunity to experience reality in a low-risk environment (Garris et al, 2002) provided effective learning opportunities as documented through the expanded list of factors named by the students during reflection exercises, positive survey responses on their engagement with the class materials, positive interview responses, and increased focus on the task as documented in the videos. As stated in the introduction, the music industry is constantly changing due to new trends and technologies and in order to build effective careers the ability to be empathetic to trends and willing to adapt is much more important than memorized facts. Hence the opportunity to create meaning and consider the impact of various actions (Malaby, 2007) that the game process provides is not easily measured in absolute facts but certainly captured in the student post reflections towards changing their habits and engagement with the industry as well as role identification during the game process.

The results suggest several options for further inquiry on using and developing games for the classroom. Given the proven effectiveness for engagement and increased learning with this specific game, options for developing the simulation over a longer period and/ or in electronic versions will be explored. One example for an extended Music Industry simulation program is the week-long immersion course offered at Columbia University in Chicago where students move through the whole cycle of creating a product and releasing it to the public. Given the importance of teacher guidance,

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
120

the development and use of an electronic simulation game should be closely monitored and guided by classroom instructors. Additional inquiry on the long term impact of learning and engagement is recommended.

Overall, this study contributes to the scholarship of teaching and learning by providing a model for developing and evaluating classroom games. The process can be applied to various fields of study and the principles of continuous evaluation and revision ensure maximum effectiveness for student engagement and motivation resulting in in-depth learning experiences.

Acknowledgments

References

Arnab, S., Brown, K., Clarke, S., Dunwell, I., Lim, T., Suttie, N., De Freitas, S. (2013). The development approach of a pedagogically-driven serious game to support relationship and sex education (RSE) within a classroom setting. *Computers and Education*, *69*, 15-30.

Barab, X., Sadler, T., Heiselt, C., Hickey, D., & Zuiker, S. (2010). Erratum to: Relating narrative inquiry and inscriptions: Supporting consequential play. *Journal of Science Education and Technology, 19*, 387-407.

Blatchford, P., Edmonds, S., & Martin, C. (2003). Class size, pupil attentiveness and peer relations. British Journal of Educational Psychology, 73(1), 15-36.

Bloom, B. S. (Ed.). (1956). Taxonomy of educational objectives, Handbook 1: Cognitive domain. New York: David McKay.

Caillois, R. (1961). *Man, play, and games*. New York: Free Press.

Chee, Y.S. & Tan, K.C.D. (2012). Becoming chemists through game-based inquiry learning: The case of Legends of Alkhemia. *Electronic Journal of e-Learning, 10*, 185-198.

Connolly, M., & Krueger, A. B. (2006). Rockonomics: The economics of popular music. *Handbook of the Economics of Art and Culture, 1*, 667-719.

Crooltall, D., Oxford, R., & Saunders, D. (1987). Towards a reconceptualization of simulation: From representation to reality. *Simulation/Games for learning, 17*(4), 147-71.

Csikszentmihalyi,M . (1990). *Flow: The psychology of optimal experience*. New York: Harper&Row.

Fenollar, P., Román, S., & Cuestas, P. J. (2007). University students' academic performance: An integrative conceptual framework and empirical analysis. British Journal of Educational Psychology, 77(4), 873-891.

Garris, R., & Ahlers, R. (2001). A game-based training model: development, application, and evaluation. Paper presented at the Interservice/Industry Training, Simulation & Education Conference, Orlando, FL.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441-467.

Glaser, B. (2003). G. & Strauss, A., L.(1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. London: Wiedenfeld and Nicholson.

Gredler, M.E. (1996). Educational games and simulations: a technology in search of a(research) paradigm. In: D.H. Jonassen (Ed.), *Handbook of research for educational communications and technology*. New York: Macmillan Library Reference USA.

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
121

Huizinga, J. (1971). Homo Ludens. New York: Beacon. de Jong, T., Martin, E., Zamarro J-M., Esquembre, F., Swaak, J., & van Joolingen, W.R.(1999). The integration of computer simulation and learning support; an example from the physics domain of collisions. *Journal of Research in Science Teaching, 36*, 597-615.

Juul, J. (2005). Half-real: *Video games between real rules and fictional worlds.* Cambridge, MA: MIT Press.

Kangas, M., Koskinen, A., & Krokfors, L. A qualitative literature review of educational games in the classroom: the teacher's pedagogical activities. *Teachers and Teaching, 23*:4, 451-470.

Leach, G. J., & Sugarman, T. S. (2005). Play to win! using games in library instruction to enhance student learning. *Research Strategies, 20*: 3, 191–203.

Malaby, T. (1999). Fateful misconceptions: Rethinking paradigms of chance among gamblers in Crete. *Social Analysis, 43*(1), 141-164.

Malaby, T. (2007). Beyond Play: A New Approach to Games. *Games and Culture,* (2). 95.

McClarty, K. L., Orr, A., Frey, P. M., Dolan, R. P., Vassileva, V., & McVay, A. (2012). A literature review of gaming in education. *Gaming in education,* 1-36.

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal, 16* (5), 455-474.

Pellegrini, A.D. (Ed.). (1995). *The future of play theory: A multidisciplinary inquiry into the contributions of Brian Sutton-Smith.* Albany, NY: State University of New York Press.

Pellegrini, A.D., & Smith, P.K. (1993). School recess: Implications for education and development. *Review of Educational Research, 63*(1), 51-67.

Provost, J.A. (1990). *Work, play, and type: Achieving balance in your life.* Palo Alto, CA: Consulting Psychologist Press.

Remus, W. and Jenner, S., 1979, Playing business games: attitudinal differences between students playing singly and as teams. *Simulation and Games, 10* (1), 75-86.

Remus, W. and Jenner, S., 1981, Playing business games: expectations and realities. *Simulation and Games, 12* (4), 480-488.

Ricci, K. E., Salas, E., & Cannon-Bowers, J. A. (1996). Do computer-based games facilitate knowledge acquisition and retention? *Military Psychology, 8*(4), 295-307.

Rieber, L.P., & Noah, D. (2008). Games, simulations, and visual metaphors in education: Antagonism between enjoyment and learning. *Educational Media International, 45*(2), 77-92. doi: 10.1080/09523980802107096.

Ruohomäki, V. (1995). Viewpoints on Learning and Education with Simulation Games, In: *Simulation Games and Learning in Production Management* (Ed: J.O. Riis), 13-25, U.K.:Chapman and Hall.

Sharritt, M. J. (2011). An open-ended, emergent approach for studying serious games. In *Serious Educational Game Assessment* (pp. 259-277). SensePublishers.

Taspinar, B., Schmidt, W., & Schuhbauer, H. (2016). Gamification in education: a board game approach to knowledge acquisition. *Procedia Computer Science, 99*, 101-116.

Taylor, T. L. (2006). *Play between worlds: Exploring online game culture.* Cambridge, MA: MIT Press.

Watson, W.R., Mong, C.J., & Harris, C.A. (2011). A case study in-class use of a video game for teaching high school history. *Computers & Education, 56,* 466-474.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., ... & Conkey, C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & gaming, 40*(2), 217-266.

Yawkey, T.D., & Pellegrini, A.D. (Eds.). (1984). Child's play: Developmental and applied. Hillsdale, NJ: Lawrence Erlbaum Associates.

Young, M. S., Robinson, S., & Alberts, P. (2009). Students pay attention! Combating the vigilance decrement to improve learning during lectures. *Active Learning in Higher Education, 10*(1), 41-55.

Teaching Developmental Theory with Interrupted Video Case Studies

Bill Anderson
Illinois State University

*Abstract***:** *This study sought to determine the usefulness of interrupted case studies, utilizing a progressive disclosure of information over time, to increase critical thinking and student learning in the study of foundational theories in the human development field. Apted's (2013) Up documentary series, consisting of video interviews over a 49-year period, was used as the interrupted study and successfully provided vicarious, but meaningful, opportunities to consistently and authentically apply course content. Participants (N = 23) were students in three sections of a graduate Human Development course where a pre-/post-test format was utilized. The effect was significant as all participant's posttest score improved* an average of 24.3%, $F(3, 19) = 3.55$, $p = .049$. *Also, coded student work indicated an increase in complex levels of thinking across the 8-week assignment, further validating post-test scores,* $t(352) = -3.172$, $p = .002$. *Evidence from student work further confirmed that an interrupted video case-study, could address limitations typically associated with case-based instruction and, more importantly, provide the critical case-study qualities needed here. Those included, telling a detailed, ambiguous, and real-life story that provided genuine context to connect theory and practice.*

*Keywords***:** *Case-based instruction, video, developmental theory*

Introduction

The lasting value of fundamental course content is largely as a means to deeper learning, successfully applying that content to problem-solving, and transferring this knowledge to future, more meaningful applications (Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010). Quite often then the most significant improvement teachers can make is to provide students more authentic observation and application experiences linked to the subject material. Yet less class time is typically available for this more meaningful, real-world learning (Mayall, 2010). When such application is impractical or impossible, other forms of doing and observing can be of value (Fink, 2003). Case-based instruction (CBI) has repeatedly been found to help generate opportunities to apply content that students might not otherwise have in business (Herreid, 2011), psychology (Mayo, 2002), engineering (Brooks, Jyothsna, & Mehmet, 2012), law (Harvard Law School, 2014), and teacher education (Harrington, 1995). CBI has also been reported to transcend the information of any single classroom (John, 2002) and promote critical thinking (Herreid, Schiller, & Herreid, 2012; Kantar, 2013, McFarlane, 2015), involving thought and action in both specific and general terms (John, 2002).

However, tangible learning outcomes with CBI have received only moderate attention (White et al., 2009). Also, and though others have found it valuable in graduate education to specifically increase critical thinking and engage in more authentic learning (e.g. Casotti, Beneski, & Knabb, 2013; Kantar, 2013; Ulanoff et al., 2009), existing research concerning CBI in graduate education is likewise limited. Also of interest to this study is Mayo's (2004) assertion that "a promising direction for case-based instruction might involve the combination of video technology and case method of teaching" (p. 144). An idea reiterated more recently by Brooks et al. (2012) and Mayall (2010) but, as with CBI in general, with little empirical work.

Literature Review

Evidence suggests that utilizing case studies that explicitly require solutions or explanations can serve as an effective means of utilizing and reinforcing explanatory theories (Egleston, 2013; Mayo, 2004; McFarlane, 2015). Teaching students to choose and appropriately apply theoretical models helps them to intentionally relate theory to practical situations and to better understand how professionals develop and articulate their ideas (Herreid, 2012; Noorminish, Mirabolghasemi, Mustaffa, Latif, & Buntat, 2013). Unfortunately, "many students do not make the important connections between and among the facts they learn in classrooms and the larger system of ideas reflected in an expert's knowledge of the discipline" (Anderson & Krathwohl, 2001, p. 42). With CBI, students can discuss their analyses, prioritize their findings, suggest solutions and possible consequences, provide evidence, and critique and refine their solutions (Harrington, 1995). During this time they also receive, process, and understand feedback.

The qualities of a good case study are many and certainly vary with teaching goals. But even those that may appear as common sense could be overlooked, such as simply being pertinent to the class and learning objectives (McFarlane, 2015). Other qualities less often considered but most relevant to this study include:

- connects theory and practice (Anderson, Bradshaw, & Banning, 2016; Penn, Currie, Hoad, & O'Brien, 2016; Prud'homme-Genereauz, 2017)
- serves as "a context for making meaning of concepts presented during instruction… thus making understanding transparent" (Ulanoff, Fingon, & Beltran, 2009, p. 125)
- tells a *focused* story and utilizes a detailed person-in-context setting where learners understood that they were watching a real person (Ruggiero, 2002)
- contains reasonable and realistic ambiguity (BU Center for Excellence and Innovation in Teaching, 2015; Ruggiero, 2002)

However, case studies are not without recognized limitations. For instance, when students are unfamiliar with CBI it may require a greater intentionality on the part of the instructor in introducing the assignment (Penn et al., 2016). Also, if the student is simply presented with a case study and provided with a set of questions, what is very likely being assessed is simply the student's ability to locate predetermined answers directly available within the case. Consequently, "students do not learn where to go to ask the appropriate questions; they learn to answer those asked by others… they learn that the answers are in front of them" (Egleston, 2013, p. 101). Furthermore, many currently available cases have student responses, instructor write-ups, and class presentations readily obtainable online. Mayo (2004) also reported that most case studies are limited in length and therefore may relate to only a few course concepts. Finally, CBI can become a largely passive activity (Herreid, 2005), perhaps especially for students who perceive the activity as busy work (Hodges, 2005) or fail to develop empathy with the characters and situation presented (McFarlane, 2015).

Still, CBI can, with time and purposeful planning, cast the learner in the more desired active role, using "course relevant ideas and explanations to frame the case problem, to analyze it, and to explain its causes… using contextual variables in a meaningful way" (Riggiero, 2002, p. 115). At that point, they are certainly practicing the role of professional (Noorminish et al., 2013), and transferring course content to real-life contexts (Penn et al., 2016; Ruggiero, 2002). Surely students learn more when they are authentically engaged in a process (Egleston, 2013) that is "not hierarchical but rather relational and even interactive" (Fink, 2003, p. 32).

The Interrupted Case Study

In 1994, Irby described a method for addressing the above limitations and moving learners to the desired more active role. With c*ase-iterative teaching*, predetermined bits of information are presented in the order in which they originally occurred and students were then asked to think aloud and advance the examination of evidence. Interpretation, serial questioning, and justification continued until all relevant information had been shared or a consensus had been reached. In addition to being more active, this method also addressed a previously mentioned limitation as the case is developed in real time with no readily available answers for students to cite or easily download (e.g. Egleston, 2013). Others have more recently described successful teaching and learning where the case is given to students in selected, organized parts (e.g. Brooks et al. 2012; Herreid, 2012; White et al., 2009). This *interrupted* case study (ICS) also uses a stepwise disclosure of information rather than revealing the entire story line at the outset, has ambiguities, but requires more time to give a necessary richness to the unfolding story (Brooks et al., 2012; Herreid, 2007). Students begin with incomplete data, create tentative hypotheses, gather more information, refine their hypotheses, predict, consider new information, and continue "to refine their thoughts and processes as additional data is received" (Brooks et al., 2012, p. 2). It is rather a progressive discovery viewed as problem-based learning over time with the learner never knowing how it will all play out and requires some speculation and risk-taking on the student's part. This format has been successful in part because it provides additional structure to the conversation, an important point for those students who do not readily engage in an unrestricted, exploratory discussion (Herreid, 2011). Additionally, the interrupted approach allows instructors multiple opportunities to pose questions, review student responses, and use those responses to address student misconceptions (e.g. Prud'homme-Genereauz, 2017).

However, when using an interrupted format, the instructor must risk no longer being the final authority, but rather a facilitator, guiding a more flexible discussion at a varying pace (Brooks et al., 2012), guiding students in "the direction of inferences and conclusions, rather than providing them with ready-made answers" (Mayo, 2004, p. 143). As "critical thinking can't be just the content of the discipline but must also be the way we go about problem-solving and asking questions" (Herreid et al., 2012, p. 21), it is the student who must lead such learning (Herreid, 2011; Kantar, 2013). Of course, students still employ textbook material and call upon personal experience (Egleston, 2013), but they are again encouraged to think aloud, essentially rehearsing professional thinking (e.g. Fink, 2003; Irby, 1994), and even to infer and imagine future consequences from the case (Anderson, Jorns, & Bivens, 2017; Herreid, 2011). Although effective teaching is always more than a simple information exchange, if the instructor is any less directive with ICS, students could possibly perceive cases as an artificially burdensome way of learning.

Purpose of this study

This study sought to examine the following points. First, to assess the usefulness of a unique *interrupted* form of CBI to increase critical thinking (e.g. Casotti et al., 2013; Kantar, 2013; Ulanoff et al., 2009). Indeed, as "factual knowledge exists at a relatively low level of abstraction" (Anderson & Krathwohl, 2001, p. 45), it seems likely that more complex thinking would yield improvement in student learning. Secondly, to determine if Apted's (2013) *Up* Series documentary could function as a fitting case study with this course material, provide Fink's (2003) meaningful vicarious opportunities to apply course content, and demonstrate the essential qualities of CBI previously mentioned.

Methods

Participant were 23 female students enrolled in three annual sections of a graduate course pertaining to human development and culture. All were first-year graduate students and most were pursuing degrees in Human Development or Child Life Specialization ($N = 18$). Other fields of study represented included one student each from: counseling, nursing, kinesiology, developmental psychology, and social work.

ICS application

The *Up* Series documentary (Apted, 2013) was used as an ICS assignment. This longitudinal series of interviews follows several individuals from 1963, at age seven, revisiting them every seven years until 2013, at age 56. Because of time constraints, four of the individuals in the film were selected by the class and professor in each class, with one or two students assigned specifically to follow each of those participants selected. Although the entire class would view each of the selected participants each week and were free to discuss each other's findings and questions as well as their own. Working with interview data at each age, students used the assumptions, concepts, and terminology of relevant developmental theory to make tentative predictions of who this child might be when they saw them at the next age. Students then continued to view the selected individuals over an eight-week period, not progressing to the next participant age until the next class meeting.

In weekly reflective essays, students anticipated growth and change in the developing participants. These essays included student's reviews of their most recent predictions, current descriptions of the assigned individuals, and their new predictions for the next seven years utilizing appropriate developmental theories and terminology. Each consecutive week, students shared their predictions for the target age with the class and then viewed the documentary participants at the new age, 14, 21, 28, 35, 42, 49, and 56. Using new information and shared insight, students refined their work, and made the next set of predictions.

Analysis

A case study format was used as it is particularly suited to examine educational process and change over time (McKinney, 2007) and to illustrating causation research (Krathwohl, 2009; Yin, 2009). However, because case studies in education typically focus on a single course or assignment results are often difficult to perceive objectively (Bishop-Clark & Dietz-Uhler, 2012; McKinney, 2007). Consequently, triangulation is more necessary in case-study research than with other designs (Krathwohl, 2009; Yin, 2003) in order to clarify the findings and result in "a better, richer, and more complete picture" (Bishop-Clark & Dietz-Uhler, 2012, p. 59). Others have reported that triangulation is also a necessary, but underused, element in studies of teaching and learning (Divan, Ludwig, Matthews, Motley, & Tomljenovic-Berube, 2017). Krathwohl's (2009) model of triangulation as a multimethod-multimeasure procedure was applied here, utilizing three independent measures, described in more detail below, to reveal any potential redundancy in the data (e.g. Divan et al., 2017; Krathwohl, 2009). Accordingly, three classes, replicated across time, were also used to examine the same learning outcome.

Pre-/Post-test

Because the low number of participants could clearly result in low statistical power, a repeated measures pre-posttest design was utilized across three equivalent replications (Bishop-Clark & Dietz-

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
126

Uhler, 2012; Wilson-Doenges, 2013). This design would also serve to moderate selection bias and better support internal validity as students are compared only to themselves (Bartsch, 2013). A pre-test consisting of 39 understanding/applying level (e.g. Anderson & Krathwohl, 2001), multiple-choice questions was administered before the first showing of *56-Up* in class. The test questions related solely to human developmental theory (Piaget, Erikson, Bronfenbrenner), cognitive development (Perry), and motivation (expectancy/value theory, motivational sequence). An identical post-test was given at the completion of the ICS assignment.

A univariate ANOVA was used to determine significant difference between pre-test scores in the three classes. Next the difference between pre- and post-test scores was determined utilizing an ANCOVA to control for the pre-test score. Finally, as recommended by Wilson-Doenges (2013), a series of paired samples *t*-tests were used to compare pre-/post- results within the three individual cohorts.

Coding

Student essays #2, at target age 14, and #8, at age 56, containing student's review of their most recent predictions, descriptions of the target individuals at the new age using appropriate theories, and their predictions for the next seven years, were coded line-by-line to determine agreement with a pre-determined pattern (e.g. Yin, 2003, 2009). Essentially, the goal was to explain relationships between data by comparing an expected theoretical pattern with an observed pattern, in this case with the pre/post results. However, pattern-matching here varied slightly from the original idea of generating an expected, conceptual pattern from existing literature (Almutairi, Gardner, & McCarthy, 2014; Yin, 2009), to utilizing an established educational pattern with Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001). It was anticipated that the coding would reveal more complex reasoning in essay #8 (e.g. Valcke, De Wever, Zhu, & Deed, 2009) and that this increase would agree with changes in the post-test results.

Descriptors for each taxonomy level, derived from the *Quick Flip Questions for the Revised Blooms Taxonomy* (2001) and listed in Table 1 below, were also used to better clarify coding concepts. This same procedure had previously worked successfully both in the pilot of the current study (Anderson et al., 2016) and later in assessing the efficacy of concept mapping in undergraduate classes (Law, Meyer, & Fall, 2016).

Table 1. Student Essay Coding Scheme

Code	Taxonomy level	Example descriptors
1	Remembering	tell, recall, relate, label, who, when, where
2	Understanding	explain, summarize, relate, rephrase, compare, contrast
3	Applying	identify, apply, utilize, organize, develop, make use of
4	Analyzing	analyze, compare, distinguish, simplify, categorize
5	Evaluating	interpret, prioritize, disprove, justify, value, conclude
6	Creating	predict, elaborate, imagine, combine, originate, propose

Intercoder reliability of two independent coders was refined over two reviews of the student essays and determined by using the intra-class correlation coefficient function of SPSS v.20 to determine a *kappa* score of reliability. A score of .80 was deemed reliable (Bishop-Clark & Dietz-Uhler, 2012). A paired samples *t*-test was used to compare initial and final essays from the three classes collectively and a series of paired samples *t*-tests to compare the three classes.

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
127

IDEA Student Evaluations

The three course sections reflected here were each evaluated by all students using the *Individual Development and Educational Assessment* instrument (IDEA, 2017), designed to specify progress on relevant teaching objectives. Results are reported on a 5-point scale as follows: 1 = No apparent progress; 2 = Slight progress; 3 = Moderate progress; 4 = Substantial progress; 5 = Exceptional progress. Each rating is presented in an unadjusted and an adjusted form. The unadjusted score, used here, does not reflect factors beyond the instructor's control that may influence student ratings (i.e. student effort not attributable to the instructor or course, difficulty not attributable to the instructor). Items deemed relevant for examination here include the following IDEA (2018) learning objectives:

- *Learning to analyze and critically evaluate ideas, arguments, and points of view*
- *Acquiring an interest in learning more by asking my own questions and seeking answers*
- *Found ways to answer my own questions*

The student rating on the above items for the three years of this study will be compared to the previous three years of IDEA ratings of the course that had not used the *Up*-Series (Apted, 2013) as an ICS utilizing a paired-samples *t*-test.

Results

Pre-/post-test

For all participants the posttest score was higher, moving from a pretest mean of 17.96/39 (*SD* – 1.00) to a posttest mean of 27.43/39 (*SD* – .874). A univariate ANOVA resulted in no statistically significant difference in pre-test scores between the three classes, $F(2, 20) = .496$, $p = .616$. Analysis with an ANCOVA revealed a significant improvement from pre- to post-test, $F(3, 19) = 3.55$, $p = .049$. A series of paired sample *t*-tests, Table 2 below, showed that the effect was significant for each cohort as well.

Table 2. Pre-/post-test scores

| Class	n	Pre-test M(SD)	Post-test M(SD)	Change	Significance
2015	7	18.86/39 (3.761)	26.43/39 (3.505)	+19.4%	$t(6) = -7.990$, $p = .000$
2016	8	18.13/39 (3.907)	27.88/39 (4.454)	+25.0%	$t(7) = -8.899$, $p = .000$
2017	8	16.88/39 (4.051)	28.00/39 (4.175)	+28.5%	$t(7) = -16.803$, $p = .000$

Coded Essays

Following coding, the intra-class correlation coefficient function of SPSS v. 20 determined a *kappa* score of inter-rater reliability of .846 for essay #2 and .818 for essay #8. Results indicated a movement toward higher level thinking on the taxonomy in each of the three classes, with a greater reliance on evaluating and creating in the final essay.

Table 3. Results of student essay coding for each class

Coding Level	2015 #2	2015 #8	2016 #2	2016 #8	2017 #2	2017 #8
Creating %	9.6	12.2	6.7	12.3	1.5	9.9
Evaluating %	9.7	12.2	8.8	24.1	15.9	18.3
Analyzing %	15.9	28.6	19.8	6.1	15.9	16
Applying %	12.4	5.5	14.1	11.8	19.5	14.3
Understanding %	17.9	8.8	27.7	19.8	27.2	15.3
Remembering %	34.5	32.7	26.6	25.9	20.0	26.2

Considering the three classes together, Table 4 also shows an overall decrease in understanding and applying responses and an increase in evaluating and creating level responses in essay #8.

Table 4. Total Coded Essay Responses
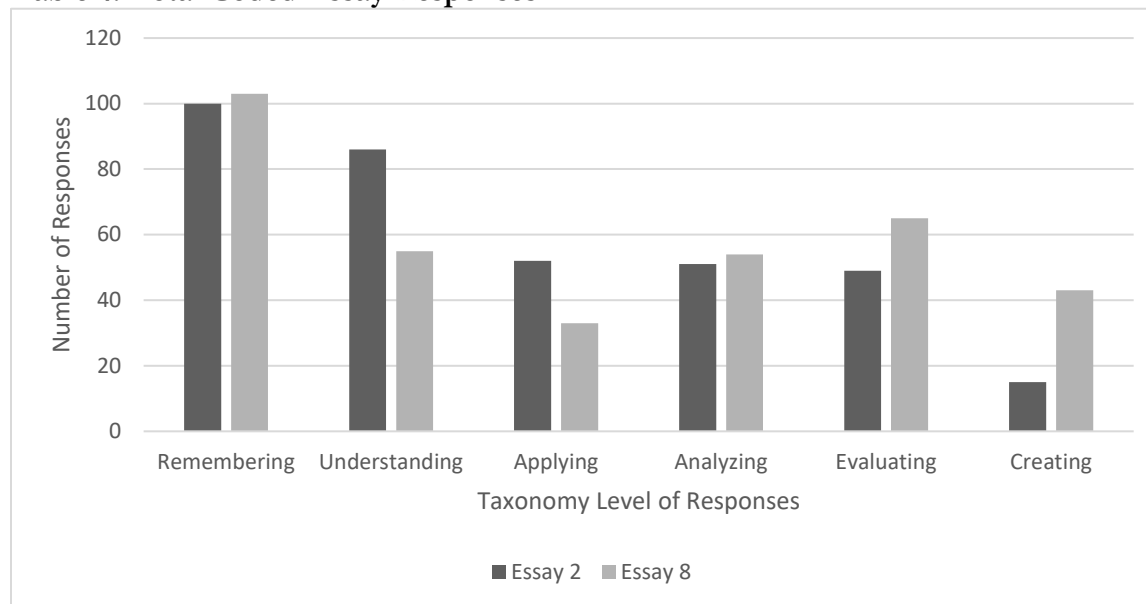
Results of a paired samples t-tests revealed that overall differences from essay 2 ($M = 2.74$, $SD = 1.548$) to the final essay ($M = 3.15$, $SD = 1.816$) were statistically significant when all three classes were combined, $t(352) = -3.172$, $p = .002$. Next, a series of paired samples t-tests were used to examine more specific changes in coding involving the total sample. Results showed that although statements at the remembering level did not change significantly, $t(2) = -.493$, $p = .671$, responses at the understanding level dropped significantly from essay 2 ($M = 24.267$, $SD = 5.52$) to essay 8 ($M = 14.633$, $SD = 3.19$), $t(2) = 8.129$, $p = .015$. Similarly, those at the applying level also decreased from essay 2 ($M = 15.333$, $SD = 3.71$) to the final essay ($M = 10.533$, $SD = 4.534$), approaching significance, $t(2) = 3.574$, $p = .070$. Statements at the analyzing level, with an essay 2 mean of 17.2 ($SD = 2.517$) and an essay 8 at 16.9 ($SD = 11.277$), showed no significant change, $t(2) = .039$, $p = .972$. Although evaluating level responses were used more in the final essay in each of the three classes, moving from $M = 11.467$ ($SD = 3.866$) to $M = 18.2$ ($SD = 5.951$), the difference was not significantly significant, $t(2) = -1.572$, $p = .257$. Responses at the creating level increased most noticeably, moving from $M =$

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
129

5.93 (*SD* = 4.104) to *M* = 11.467 (S*D* = 1.358) and approached statistical significance, *t*(2) = -3.304, *p* = .081.

IDEA student ratings

Each of the previously identified IDEA (2018) learning objectives were compared to the previous three years of IDEA ratings for the course that had not used the *Up-Series* (Apted, 2013) as an ICS. Most showed a slight improvement.

- *Learning to analyze and critically evaluate ideas, arguments, and points of view*, improved from *M* = 4.67/5.0 (*SD* = .153) to *M* = 4.83/5.0 (*SD* =.208). Results were significant, *t*(2) = -5.00, *p*. = .038
- *Acquiring an interest in learning more by asking my own questions and seeking answers*, improved from *M* = 4.37/5.0 (*SD* = .231) to *M* = 4.63/5.0 (*SD* =.305), with results approaching significance *t*(2) = -4.00, *p*. = .057
- *Found ways to answer my own questions* showed no change at 4.84/5.0

Discussion

Results indicated an increase in student learning for all participants with an average increase in post-test performance of 24.3%. Coding of student essays with Anderson and Krathwohl's (2001) revised taxonomy supported the posttest improvement with an indicated student movement toward a more complex thinking. More specifically, students utilized *understanding* (explain, summarize, rephrase) and *applying* (identify, utilize, organize) less and increased their use of creating (predict, elaborate, propose). It seems likely that both of these results are, at least in part, the consequence of students having opportunities to actually apply theory, evaluate their application, process feedback, and refine their position over time. Therefore, these findings support existing research that suggests utilizing ICS that require solutions or explanations can serve as an effective means of applying and reinforcing explanatory theories (Egleston, 2013; Mayo, 2004; McFarlane, 2015; Prud'homme-Genereaux, 2017). This is further evidenced in the following student responses:

> Suzy listened to her parents and did what was expected of a child in her environment. Now that she is an adult… Foreclosure? Isolation?

> Nicolas demonstrated a form of thinking between relativism and commitment… couldn't that relate to the person's developmental niche?

> Nicholas is already showing Perry's commitment… [he] states that the individual looks at priorities and commits fully to one; however if convinced, may be inclined to change his or her mind.

Results here also supported the idea that teaching students to choose and appropriately apply theoretical models helps to intentionally relate theory to practical situations and leads to better understanding how professionals in the human development and family science field advance and communicate their ideas (e.g. Herreid, 2012; Noorminish et al., 2013). Related student responses included:

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
130

She has such short-term goals that she is not attempting to prioritize thinking. I found this interesting because, just as Piaget's stages (formal thinking) were linked to Erikson's stages (identity), I now think that Perry's reflective judgment should be as well.

I am now linking [expectancy*value] motivation with the stages of psychosocial development and the developmental niche.

Overall, my work evolved from basic understanding and knowledge, to the creation of new ideas, finishing with unanswered questions.

Although findings from the IDEA ratings comparisons were less robust compared with those from the pre-/post-test and coding, results indicated a subtle improvement in the area of considering and evaluating other's perspectives, certainly critical to the ICS assignment. Perhaps this was related to the evaluating/creating level descriptors previously mentioned. Support from student observations included:

I am better able to understand that I must examine an individual's culture in order to fully understand them as a person. I am a little surprised as I look back at some of my earlier work. I was so confident but using so little information.

A second trend I noticed while reviewing my work from this semester involved considering between and within group differences.

The second goal was to ascertain if the *Up*-Series (Apted, 2013) could meet the needs of teaching developmental theory as a video case study (e.g. Brooks, et al., 2012; Mayall, 2010; Mayo, 2004). Results here indicate that the video interviews served successfully in this role, and also illustrated that the ICS met several standards of CBI. Specifically, it was pertinent to the class learning objectives (McFarlane, 2015), encouraged retention (Penn et al., 2016), and consistently linked theory and practice (Penn, Currie, Hoad, & O'Brien, 2016; Prud'homme-Genereauz, 2017). The series addressed a number of other essential CBI qualities as well.

First, recall that Ulanoff et al. (2009) described cases as teaching tools that make learning more accessible by serving as frameworks for student meaning-making. In this case, following the participants (Apted, 2013) over a 49 year period provided authentic context for understanding (e.g. Ulanoff et al., 2009) developmental theory. Student remarks included:

Commitment in relativism is recognizing choices, accepting responsibility and consequences, and willingness to accept others' right to their choices. Suzy shows full recognition that choices restrict some choices and open the way to others. She may continue to reaffirm or reject old beliefs; either way, the decision is based on a conscious consideration of alternatives. She is no dualist!

I began the course being able to criticize others' ideas and research, provide specific evidence regarding why I made certain predictions, and be curious about and interested in the experience of others in a different culture. I ended the course feeling confident in my ability to seek and find an answer to my own questions.

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
131

Next, this person-in-context setting element of CBI must communicate a detailed and focused human story (Ruggiero (2002), engaging and meaningful to students (Herreid 2007; Mayo, 2004). Apted's (2013) interviews allowed students to discover and understand "that, for someone else, this dilemma is very real" (Ruggiero, 2002, p. 120) and confirmed Herreid's (2007) statement, "there is no better way to understand a situation and to gain empathy for the characters than to hear them speak in their own voices" (p. 46). Students here regularly responded with interest, insight, and sometimes concern.

> I had worried about Bruce every week, despite his continued success as a teacher. But last week, seeing his wedding (finally) I actually cried a little in class, but I wasn't the only one.

> Neil (at 56) said that he was absolutely sure that his faith helped him through difficult times… [and] again shows signs of, I believe, generativity. But then we have been describing Neil this way since he was 28 haven't we?

> I would like to predict that his life becomes a little more stable, but I'm not sure if that is a true prediction or just my hope.

Lastly, CBI is made more authentic with the inclusion of realistic ambiguity and uncertainty (BU Center for Excellence and Innovation in Teaching, 2015; Ruggiero, 2002). With the presence of this quality, students "never know how it will all come out..." (Herreid, 2007, p. 48). Evidence from students included:

> The more I watch these episodes, the more I realize that anything can happen in a seven year span.

> Was Neil simply more reflective at 14, as I had thought, or did I actually miss something that would have let me accurately predict Neil at 21? Was I thinking theory or not seeing past what I was hoping for Neil.

> I might be thinking this way because of the way I feel about my parents… assuming his experience is reflective of mine. I think that the distance Neil has with his family might be for a reason that we have yet to see in his interviews. He seemed less poetic… I got the impression that he might be using a drug therapy to manage... I also didn't really see the "sparkle of the 7-year old eyes." It felt like meeting a whole new person… I hope he remains in good health, but I am worried.

Conclusion

The use of the interrupted *video* case study format was successful in satisfying the first goal of this study, resulting in increased critical thinking (Herreid, 2012; Brooks et al., 2012; White et al., 2009) and better learning and retention (McFarlane, 2015; Penn et al., 2016). Indeed, the interrupted format encouraged decision making and explanation building (e.g. Ruggiero, 2002), and guided students toward progressive, step-wise inferences, rather than singular, ready-made conclusions (Mayo, 2004). The independent and varied methods used here, considered as a multimethod-multimeasure procedure (Krathwohl, 2009), resulted in a clear, triangulated support (e.g. Krathwohl, 2009; Yin, 2003, 2014) of student learning anchored in the ICS curriculum. Results also support the idea that

Journal of the Scholarship of Teaching and Learning, Vol. 19, No. 5, December 2019.
josotl.indiana.edu
132

CBI is beneficial in graduate education by specifically increasing critical thinking (e.g. Casotti et al., 2013; Kantar, 2013; Ulanoff et al., 2009).

Concerning the second goal, Apted's (2013) *Up*-Series effectively provided authentic opportunities (e.g. Fink, 2003) to apply theory (e.g. Prud'homme-Genereauz, 2017), and met all necessary requirements for an effectual CBI. Unlike cases that are typically more limited in length and scope, relating to only some course concepts (e.g. Mayo, 2004), Apted's (2013) documentary covers 49 years of observations and interviews of people of varied gender, ethnicity, and socio-economic status. Furthermore, the series is *interrupted* by design and easily provided a consistent, unfolding story of participant's lives. The resulting inevitability of reevaluating and reorganizing previous learning and information appears to have allowed students to "disentangle the normative (what ought to be) and the operative (what is)" and "travel confidently between theory and practice" (John, 2002, p. 337) resulting in greater use of higher level thinking. In the process, students grew accustomed to the necessary ambiguity (e.g. BU Center for Excellence and Innovation in Teaching, 2015) and progressed in their understanding of the application and implicit flexibility of theory. Lastly, and although Egleston (2013) reported that students have expressed some dislike for the less structured format that might accompany case-based learning, that was not an issue in this current study. At the conclusion of the assignment, and in agreement with Mayo (2002), students here were very favorable concerning CBI, as well as the interrupted format.

Future Research

Certainly the understanding and application of ICS will benefit not only from the necessary replications of the current study but also from consideration by other fields of study. The *Up* Series (Apted, 2013) used here could feasibly be used similarly in sociology or social psychology. For instance, a theoretical emphasis on attribution theory (e.g. Prud'homme-Genereauz, 2017), as opposed to developmental theory, would allow students to apply concepts such as locus of control or the persistence of social stereotypes, or to consider culture, poverty, and tradition from an ecological theory perspective.

Following an ICS format using intentionally selected books could also provide the desired ambiguous stepwise disclosure (Brooks, et al., 2012), the potential of empathic reaction (McFarlane, 2015), and a real-person-in-context setting (Ruggiero, 2002). Possibilities here might include Alvord's (2000) *The Scalpel and the Silver Bear* to explore the process of culture shock or Pausch's (2008) *The Last Lecture* as an ICS in discussions of bereavement of grief. The possibilities are certainly not limitless but they are indeed many, and would further the effort of seeking best practice with ICS.

<div align="center">

References

</div>

Almutairi, A.F., Gardner, G.E., & McCarthy, A. (2014). Practical guidance for the use of pattern-matching technique in case-study research: A case presentation. *Nursing and Health Sciences, 16*, 239-244. doi:10.1111/nhs.12096

Ambrose, S.A., Bridges, M.W., DiPietro, M., Lovett, M.C., & Norman, M.K. (2010). *How learning works: Seven research-based principles for smart teaching.* San Francisco: Jossey-Bass.

Anderson, B. (2017, October). Theoretical pattern-matching. *The SoTL Advocate*, Web log post at https://illinoisstateuniversitysotl.wordpress.com/2017/10/30/theoretical-pattern-matching-in-sotl-sotl-methods-series-1/

Anderson, B., Bradshaw, S., & Banning, J. (2016). Using interrupted video case studies to teach developmental theory: A pilot study. *Gauisus, 4.* Retrieved from http://gauisus.weebly.

com/anderson-et-al---using-interrupted-video-case-studies-to-teach-developmental-theory-a-pilot-study.html

Anderson, B., Jorns, G.S., & Bivens, L. (2017). Using interrupted case studies to address case study limitations. Paper presented to annual Teaching and Learning Symposium of the Center of Teaching and Learning Technology. Normal, IL.

Anderson, L.W., & Krathwohl, D.R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Blooms taxonomy of educational objectives.* London: Longman, Inc.

Apted, M. (Director). (2013). *The up series.* London: First Runs and Features.

Bartsch, R.A. (2013a). Designing SoTL studies – Part 1: Validity. In R.A.R. Gurung & J.H. Wilson (Eds.), *Doing the scholarship of teaching and learning: Measuring systematic changes to teaching and improvements in learning.* (pp. 17-34). San Francisco: Jossey-Bass.

Bishop-Clark, C. & Dietz-Uhler, B. (2012). *Engaging in the scholarship of teaching and learning: A guide to the process and how to develop a project from start to finish.* Hemdon, VA: Stylus Publishing.

Brooks, R.M., Jyothsna, K.S., & Mehmet, C. (2012). Interrupted case method for teaching ethics and transportation engineering and systems management course. Retrieved from https://peer.asee.org/interrupted-case-method-for-teaching-ethics-in-transportation-engineering-and-systems-management-course.

BU Center for Excellence and Innovation in Teaching. (2015). *Using case studies to teach.* Retrieved from http://www.bu.edu/ceit/teaching-resources/using-case-studies-to-teach/

Casotti, G. Beneski, J. T., & Knabb, M. T. (2013). Teaching physiology online: Successful use of case studies in a graduate course. *Advances in Physiology Education, 37*, 65-69. doi:10.1152/advan.00159.2012.

Divan, A., Ludwig, L., Matthews, K., Motley, P., & Tomljenovic-Berube, A. (2017). Survey of research approaches utilized in the scholarship of teaching and learning publications. *Teaching & Learning Inquiry*, 5(2). doi:10.20343/teachlearninqu.5.2.3.

Egleston, D.O. (2013). The interactive, progressive case study. *Business Education Innovation Journal, 5*(1), 101-104. Retrieved from http://beijournal.com/images/V5N1_Abstracts.pdf

Fink, L.D. (2003). *Creating significant learning experiences: An integrated approach to designing college courses.* San Francisco: John Wiley & Sons.

Harrington, H.L. (1995). Fostering reasoned decisions: Case-based pedagogy and the professional development of teachers. *Teaching & Teacher Education, 11*(3), 203-214. doi:10.1016/0742-051X(94)00027-4

Harvard Law School. (2014). The case teaching method. Retrieved from http://casestudies.law.harvard.edu/the-case-study-teaching-method/.

Herreid, C.F. (2005). The interrupted case method. *Journal of College Science Teaching, 35*, 4-5. Retrieved from http://sciencecases.lib.buffalo.edu/cs/pdfs/Interrupted%20Case%20Method-XXXV-2.pdf

Herreid, C.F. (2007). What makes a good case? In C.F. Herreid (Ed.), *Start with the story: The case study method of teaching college science* (pp. 45-48). Arlington: NSTA Press.

Herreid, C.F. (2011). Case study teaching. In W. Buskist, and J.E. Groccia (Eds.), *Evidence-based teaching* (pp. 31-40). Hoboken, NJ: Wiley.

Herreid, C.F. (2012). Can case studies be used to teach critical thinking? In C.F Herreid, N.A. Schiller, and K.F. Herreid (Eds.), *Science stories: Using case studies to teach critical thinking* (pp. 21-24). Arlington: NSTA Press.

Herreid, C.F., Schiller, N.A., & Herreid, K.F. (2012). *Science stories: Using case studies to teach critical thinking.* Arlington, VA. National Science Teachers Assoc.

Hodges, L.C. (2005). From problem-based learning to interrupted lecture: Using case-based teaching in different class formats. *Biochemistry and Molecular Biology Education, 33*(2), 101-104. https://doi.org/10.1002/bmb.2005.494033022446

IDEA. (2017). Student ratings of instruction. Retrieved from http://www.ideaedu.org/Services/ Services-to-Improve-Teaching-and-Learning/Student-Ratings-of-Instruction

IDEA (2018). IDEA notes on instruction. Retrieved from https://www.ideaedu.org/Resources-Events/Teaching-Learning-Resources/IDEA-POD-Notes/Notes-on-Instruction

Irby, D.M. (1994). Three exemplary models of case-based teaching. *Academic Medicine, 69* (12), 947-953.

John, P.D. (2002). The teacher educator's experience: Case studies of practical professional knowledge. *Teaching and Teacher Education, 18*(3), 323-341. Retrieved from http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.297812

Kantar, L.D. (2013). Demystifying instructional innovation: The case of teaching with case studies. *Journal of the Scholarship of Teaching and Learning, 13*(2), 101-115. Retrieved from https://eric.ed.gov/?id=EJ1011686

Krathwohl, D.R. (2009) *Methods of educational and social science research: The logic of methods* (3rd ed.).Long Grove, IL: Waveland Press.

Law, D.D., Meyer, S., & Fall, L. (2016). Using family science concept maps to gain higher order student learning outcomes. *Family Science Review, 21*(2), 116-134. Retrieved from http://www.familyscienceassociation.org/sites/default/files/6%20-%20LAW%20MEYER%20FALL.pdf

Mayall, H.J. (2010). Integrating video case studies into a literacy methods course. *International Journal of Instructional Media, 37*(1), 33-41. Retrieved from https://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=34&sid=85271328-abba-4e6e-a2cb-87d6724b46f9%40sessionmgr4009

Mayo, J.A. (2002). Case-based instruction: A technique for increasing conceptual application in introductory psychology. *Journal of Constructivist Psychology, 15*, 65-74.

Mayo, J.A. (2004). Using case-based instruction to bridge the gap between theory and practice in psychology of adjustment. *Journal of Constructivist Psychology, 17*, 137-146.

McFarlane, D.A. (2015). Guidelines for using case studies in the teaching-learning process. *College Quarterly, 18*(1). Retrieved from https://eric.ed.gov/?id=EJ1070008

McKinney, K. (2007). *Enhancing learning through the scholarship of teaching and learning*. Bolton, MS: Anker Publishing.

Noorminish, A.I., Mirabloghasemi, M., Mustaffa, N.H., Latif, M.S., & Buntat, Y. (2013). Student perception of using case study as a teaching tool. *Social and Behavioral Sciences, 93*, 2200-2204. https://doi.org/10.1016/j.sbspro.2013.10.187

Penn, M.L., Currie, C.S.M., Hoad, K.A., & O'Brien, F.A. (2016). The use of case studies in OR teaching. *Higher Education Pedagogies, 1*(1), 16-25. http://dx.doi.org/10.1080/23752696.2015.1134201

Perry, W.G. (1999). *Forms of ethical and intellectual development in the college years: A scheme*. San Francisco: Jossey-Bass.

Prud'homme-Genereaux, A. (2017). Formulating questions that address student misconceptions in a case study. *Journal of College Science Teaching, 46*(4), 54-60. Retrieved from http://www.nsta.org/store/product_detail.aspx?id=10.2505/4/jcst17_046_04_54 *Quick Flip Questions for the Revised Blooms Taxonomy* (2001). Madison, WI: Edupress.

Ulanoff, S.H., Fingon, J.C., & Beltran, D. (2009). Using case studies to assess graduate's knowledge and skills in a graduate reading program. *Teacher Education Quarterly, 36*(2), 125-142. Retrieved from https://archive.org/details/ERIC_EJ857480

Valcke, M., De Wever, B., Zhu, C., & Deed, C. (2009). Supporting active cognitive processing in collaborative groups: The potential of Bloom's taxonomy as a labeling tool. *Internet and Higher Education, 12,* 165-172. doi:10.1016/j.iheduc.2009.08.003

White, T.K., Whitaker, P., et al. (2009). The use of interrupted case studies to enhance critical thinking skills in biology. *Journal of Microbiology & Biology Education, 10*(1), 25-31. doi:10.1128/jmbe.v10i1.96

Wilson-Doenges, G. (2013). Statistical models for analyzing learning data. In R.A.R Gurung & J.H. Wilson (Eds.), *Doing the scholarship of teaching and learning: Measuring systematic changes to teaching and improvements in learning.* (pp. 49-58). San Francisco: Jossey Bass.

Yin, R.K. (2003). *Applications of case study research* (2nd ed.). London: Sage.

Yin, R.K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.

Yin, R.K. (2014). *Case study research: Design and methods* (5th ed.). London: Sage.

Contact Info for the Journal