

# From Comprehensive to Singular: A Latent Class Analysis of College Teaching Practices

Corbin M. Campbell<sup>1</sup> · Alberto F. Cabrera<sup>2</sup> · Jessica Ostrow Michel<sup>1</sup> · Shikha Patel<sup>1</sup>

Received: 10 December 2015 / Published online: 5 October 2016  
© Springer Science+Business Media New York 2016

**Abstract** While decades of research on college teaching has investigated several forms of classroom practices, much of this research approaches teaching as falling into mutually exclusive paradigms (e.g., active learning vs. lecturing). This paper enters inside the college classroom using external raters to understand patterns of pedagogical practices embedded in heterogeneous groups of courses. The study used quantitative observation and draws on data from a multi-institutional study of 587 courses across nine institutions to understand the patterns of teaching practices within courses. Latent class analyses demonstrated that there were five patterns of seven course practices that cluster around active learning, lecturing, and cognitively responsive practices: Comprehensive, Traditional Lecture, Active Lecture, Integrative Discussion, and Active Only.

**Keywords** College teaching · Latent class analysis · Active learning · Cognitive responsiveness · Observation

## Introduction

Understanding good college teaching practices has become increasingly important with continued public affronts on the quality of higher education (Arum and Roksa 2011; Carey 2014; Guttenplan 2014; Kaminer 2013). With rising costs of college and diminishing public trust in higher education, the public media has questioned the educational process inside college and university classrooms (Carey 2014). Responding to these public concerns, policy-makers have considered the importance of college teaching in preparing a 21st century workforce and democracy. In one example, the call for better teaching in the

---

✉ Corbin M. Campbell  
campbell2@tc.columbia.edu

<sup>1</sup> Teachers College, Columbia University, 525 W. 120th St, Box 101, New York, NY 10027, USA

<sup>2</sup> University of Maryland, College Park, MD, USA

Science, Technology, Math, and Engineering (STEM) fields has reached the level of a national imperative, as policy-makers consider improvements in college teaching as central to broadening student participation in STEM fields and staying at the forefront of scientific and technological advances (Henderson et al. 2011). Henderson et al. engaged in a meta-analysis of the literature on reforms in STEM college teaching and found that researchers across three fields have taken up this call to action, including disciplinary-based education researchers, faculty development researchers, and higher education researchers. Beyond STEM, considering the cross-disciplinary work on college teaching, we would add the learning sciences to this list of fields advancing the understanding of college teaching strategies largely from a cognitive perspective. Taken together, these fields have mounted a great deal of evidence about effective college teaching, yet the changes that professors have made in classroom practices have not been as substantial or pervasive as could be possible given the level of evidence of efficacious practices. Beach et al. (2012) comment on the state of moving evidence to practice in college teaching:

“But despite having significant knowledge of effective practices, along with curricular and pedagogical resources, efforts to transform introductory sequences have met with only modest success. While we may know what to do, we do not know how to enact and sustain these reforms at scale.” (p. 53).

Studies of trends in faculty teaching practices across disciplines provide context to Beach et al.'s (2012) concerns that the evidence for what works in college teaching practices has only been made actionable by a proportion of faculty. For example, results from the 2013–2014 national survey of faculty from the Higher Education Research Institute, a study that was across disciplines, indicated that less than half of faculty assigned group projects and 56 % reported using student inquiry to drive learning (Eagan et al. 2014). One promising finding was that 83 % of faculty reported incorporating the use of class discussions, but this appears not to be a dominant practice during the class period given that 51 % of faculty reported using extensive lecturing in all or most of their classes. Additionally, these results are self-reported by faculty whose responses may suffer from social desirability bias, causing upward estimates of active learning given that it is a well-known evidence-based practice. While substantive scholarship has focused on how faculty perceive their teaching at large (e.g., DeAngelo et al. 2009; Eagan et al. 2014; Umbach and Wawrzynski 2005) and how students perceive college teaching (e.g., NSSE 2015; Reason et al. 2006), little research examines the pervasiveness in evidence-based college teaching practices as embedded in courses. The few multi-institutional studies that do examine specific classroom practices at the course level were largely based on self-reports of students or faculty (e.g., Cabrera et al. 2001; Volkwein et al. 2007). A notable exception is Hora and Ferrare's (2014) study, which relied on external observers to understand patterns of teaching practices using networks analysis, but the study had a small sample (58 instructors across three institutions) and viewed classes individually using a variable-centric approach. Interestingly, though, single institution studies have had more methodological diversity in the study of active learning practices. For example, Lammers and Murphy (2002) conducted an observational study of 58 college classrooms at a U.S. public university, and found that, on average, courses were instructor driven (i.e. no students actively engaged in course material) almost 50 % of the total class time. Yet, these single institution studies do not yield a broader perspective for the field. In order to understand pervasive trends in evidence-based college teaching, practices should be studied across multiple institutions, but understood as situated within the contexts of classrooms. The purpose of this multi-institutional study was to enter inside the college classroom using

external observers to understand teaching practices at the course level by discerning patterns of pedagogical practices embedded in heterogeneous groups of courses. In essence, do courses display patterns of teaching practices in a systematic way as appraised by individuals other than instructors or students themselves?

## Active Learning

Although higher education scholars have considered many ways to understand college teaching, active learning has been a central focus in the literature over several decades (Carroll 1963; Freeman et al. 2014; McKeachie and Kulik 1975). Active learning can be understood as pedagogical practices that engage students in their learning process, such as class discussions, group work, instructor questions, problem-based learning, and case studies (Arends and Castle 1991; Astin 1993; Braxton et al. 2000; Carini et al. 2006; McKeachie and Kulik 1975; Pascarella and Terenzini 2005; Rothkopf 1973; Van Der Meij 1994). The active learning paradigm grew out of the movement to transition from teacher-centered to learner-centered teaching practices. According to Robertson (2005), in learner-centered courses, “teachers construe themselves to be facilitators of student learning as opposed to teacher-centered teaching where teachers see themselves as disseminators or imparters of knowledge” (p. 181). In active learning, by engaging students in the classroom, instructors allow students to co-construct their learning experience around course ideas. In McKeachie and Kulik’s (1975) review of effective college teaching strategies, they contended that “since 1970...the importance of student participation and interpersonal interaction emerges more and more clearly” (p. 199). Rooted in this movement, the past two decades of most national survey studies of college teaching practices have focused on measuring the extent to which instructors interact with and engage their students (DeAngelo et al. 2009; Kuh 2008). For example, Koljatic and Kuh (2001) examined 14 years of data from the College Student Experience Questionnaire, and found that the frequency of student involvement in active learning did not change over time.

In addition to studies exploring teachers’ facilitation of active learning practices, several studies have focused on the effects of active learning practices. These studies contribute an important voice to the conversation because students’ classroom perceptions are heavily shaped by how faculty members structure class time, as active learning is associated with greater student self-reported educational gains (Laird et al. 2008; Kuh 2008; Pascarella and Terenzini 2005). This can be seen in Braxton et al.’s (2000) study, which found that exposure to active learning practices, such as class discussion, influenced students’ institutional commitment and persistence decisions. In another example, Cabrera et al. (2001) studied survey responses from more than 1200 students in engineering courses across seven schools and found that collaborative learning and instructor interaction contributed to important student outcomes, such as group skills, problem solving skills, and occupational awareness. However, each of these studies treated teaching practices as individual variables and did not examine whether there were different patterns of courses with collective teaching practices.

The results of active learning studies from discipline-based education researchers in the STEM field move beyond student and faculty self-reports to find promise in a broad base of experimental studies. Freeman et al. (2014) conducted a meta-analysis of 225 studies of STEM undergraduate courses that compared student performance in courses with at least some active learning to courses with traditional lecturing alone. They found dramatic

results, with an effect size of 0.47, indicating that students in courses with some component of active learning scored a half a standard deviation higher when compared with students in standard lecture courses. While these results may seem conclusive, other scholars questioned whether a meta-analysis of these binary buckets (lecturing vs. active learning) was an adequate picture of teaching practices. As such, in Hora's (2014) published response to Freeman et al.'s (2014) study, he discerns that, lecturing is not a monolithic category and that faculty can use a variety of different strategies while also maintaining "continuous exposition by the teacher," which was the condition for "lecturing" in the Freeman et al. study. As such, Hora (2014) states, "skepticism is therefore warranted regarding the assumption that all of the 225 studies were measuring (or controlling for) the same type of instruction" (p. 3024). Following in this vein, the body of literature on active learning has been critiqued for using theoretical tools that are too blunt to understand the complex dynamics within a college classroom (Hora and Ferrare 2014). Teaching is a complex activity that is influenced by different behavioral factors, such as clarity, interaction, organization, enthusiasm, and feedback in a classroom (Feldman 1997; Marsh and Roche 1993). Using broad categories such as "active" or "passive" may obscure important subtleties in the teaching process, such as the interaction between multiple forms of pedagogy and the disciplinary context (Hora and Ferrare 2014; Major and Palmer 2006; Oleson and Hora 2014; Shulman 2004a).

## Developing Theory

Frameworks on college teaching culled from the learning sciences or socio-cultural studies examine the subject matter and subtler practices that may shape student learning. One example that explores such subtleties was Hora and Ferrare's (2013) study examining class observation and faculty interview data from 57 faculty across three research institutions. They found that there were systems of practice embedded within disciplines and instructors that indicated configurations of teaching methods during a class period that exhibited five dimensions of teaching practices (teaching methods, pedagogical strategies, student–teacher interactions, cognitive engagement, instructional technology). In another example, Neumann's (2014) qualitative study of a college philosophy class in a high diversity institution found that subject matter and disciplinary ideas, teacher's framing of the material, learner's prior experiences, and course context all shaped the learning process. While this qualitative study highlights the complexity of the teaching and learning process and draws in a cognitive element that was less visible in other frameworks, this study was not able to understand the breadth of this cognitively responsive paradigm across courses. Other qualitative studies of individual courses or small groups of courses (e.g., Lattuca et al. 2004) also focus on the specific disciplinary contexts that shape faculty practices.

In each of these studies, the developing teaching theory points to the use of varied teaching techniques in specific educational spaces, contexts, conditions, and for specific students. In essence, an expert instructor knows what specific core ideas for the subject matter can be taught by using lecture—and this lecture may look different, taking on various rhetorical forms, at different times during the semester with different students (Bransford et al. 2000). Perhaps instead of two binary buckets of lecturing and active learning, teaching would look more like a web weaving between content, pedagogical technique, cognitive engagement, and course context. Hora and Ferrare (2014) demonstrate this phenomenon through network analysis of instructors' teaching tools. Neumann's (2014)

work builds off the larger learning science literature base that Bransford et al. (2000) summarized in a report by the National Academies, demonstrating the important role of certain teaching practices in students' learning. For example, teaching in depth for a small number of core ideas for the course is more beneficial for learning for transfer when compared to teaching for breadth across a wide variety of ideas (Bransford et al. 2000). Additionally, using a framework to organize and sequence the ideas in a way that is specific to the discipline has been shown to aid recall transfer. Metacognition and reflection about how a student is learning—what a student knows and does not know—has also been connected to learning. In each of these examples of practices, the efficacious practices do not focus on the frequency of interaction (as does active learning), but a consideration of the context of the learner, the teacher, and the course content (Neumann and Campbell 2016). Understanding these complex dynamics that take shape between faculty, students, course content, and context is often learned through many years of practice, in several facets of faculty lived experiences, and in specific developmental experiences (Major and Palmer 2006; Neumann 2014; Oleson and Hora 2014; Shulman 2004a).

## Methods for Multi-Institutional Study of College Teaching Practices

Self-report surveys have been the prevalent means for studying college teaching across disciplines in multi-institutional studies. Faculty surveys are a commonly used instrument to study faculty teaching for a variety of reasons, such as improvement or evaluation of teaching practices. Based on a study by Carnegie Foundation for the Advancement of Teaching, approximately 82 % of the universities consider faculty self-reported data to review classroom teaching (Magner 1997). Additionally, national surveys such as the Higher Education Research Institute Faculty Survey (HERI Faculty Survey) and the Faculty Survey of Student Engagement (FSSE) use questionnaires specific to pedagogical practices to determine the perception of faculty about their own teaching practices. The HERI Faculty Survey asks faculty to report their teaching practices across classes, whereas FSSE asks faculty to report class time allocated to particular teaching methods that are linked with high levels of learning and development. As an example of the pervasiveness of the use of faculty survey to measure college teaching nationally, the most recent HERI faculty survey report (2013–2014) was based on 16,112 full-time college and university faculty members at 269 four-year colleges and universities nationwide. Another often cited method for studying faculty-teaching practices is through student self-report surveys. In one prominent example, the 2015 NSSE instrument asked approximately 315,000 first year and senior students across 541 universities about the teaching practices in their courses.

These national surveys of college teaching are useful toward understanding how students perceive their college course experiences and how faculty understand their teaching practice. Yet, the view that these methods yields is narrow in three primary ways. First, student surveys may not be able to capture developing theories that examine subtle dynamics that take shape between students, course content, and context. Students are not often equipped with adult learning theory or theory from the learning sciences that would be important when answering questions based on rich theory. Secondly, most multi-institutional studies of college teaching ask respondents to aggregate their experiences over the past year. This may reflect a general sense of student or faculty experiences at-large, but aggregating teaching experiences over the past year does not give evidence of the specific teaching practices that are important given certain course content with specific

students as embedded within a course. This is problematic when theories from the learning sciences and socio-cultural studies of teaching and learning demonstrate that pedagogical choices are made at the course level, with specific content for specific students. Finally, both faculty and students have a substantial stake on what takes place in the classroom, which may influence the perspective that self-reports yield. For example, the results of faculty self-reporting of their teaching practices may not correlate with the results of students' perception of teaching in the classroom (Centra 1973; Feldman 1989a). Perhaps an outside observer who does not have a stake in the course could offer a unique perspective.

As demonstrated by qualitative scholars of college teaching and learning, theory on college teaching points to context and complexity, and therefore, the field is in need of a nuanced window into the college teaching and learning experience. Substantive scholarship has focused on self-reports of teaching behaviors by faculty (e.g., Umbach and Wawrzynski 2005; Volkwein et al. 2007) and self-reported educational engagement by students (e.g., Cabrera et al. 2001; Reason et al. 2006). Largely missing is scholarship that unpacks this issue from inside the classroom and across multiple institutions. Studies using a quantitative observational protocol may provide useful supplementary insights into classroom practices in a way that mirrors the complexity of recent theories, but with a view that aggregates across courses. Additionally, while current studies contribute to the field's understanding of what teaching practices faculty enact (for example, what proportion of faculty use active learning and what proportion use lecture), these studies treat the occurrence of teaching behaviors individually rather than understanding that a heterogeneous population of teachers likely favors one *set of teaching practices* over another. New analytical tools enable us to understand how patterns of teaching behaviors are grouped together among certain clusters of courses. Therefore, the purpose of this study was to address the following research questions using quantitative observational data and latent class analysis (LCA):

1. Are there classes of courses that cluster according to their pedagogies (class discussions, class activities, student questions, lecture, subject-matter expertise, prior knowledge, supporting changing views)?
2. If so, what set of pedagogies is each class of courses more prone to enact?

## Theoretical Framework

The theoretical framework for this study considers two distinct conceptualizations of college teaching: the landmark theory of active learning (Carroll 1963; McKeachie and Kulik 1975) and the emerging conceptualization of cognitively responsive teaching (Neumann 2014) that rests on several prominent theories from the learning sciences (Bransford et al. 2000; Shulman 2004a, b). The conceptualization of "active learning" was developed the 1960s, and it grew up in a context of two simultaneous movements: "student-centered" learning (McKeachie and Kulik 1975) and "time on task" (Carroll 1963). Since that time, there have been a broad range of pedagogies that have been considered as active learning strategies in the literature base including, but not limited to, class discussions, group work, problem-based learning, studio-based learning, experiential education, and case studies (Arends and Castle 1991; Astin 1993; Braxton et al. 2000; Carini et al. 2006; McKeachie and Kulik 1975; Pascarella and Terenzini 2005; Rothkopf 1973; Van

Der Meij 1994). These active learning techniques both focus students' attention on course material (i.e. time on task) and create opportunities to bring forth the student's context and learning needs (i.e. student-centered learning). The basis for this theory is that students learn best by engaging in course material through interacting with peers and instructor and actively engaging in the course content.

The second framework we used, namely Neumann's (2014) framework on cognitively responsive teaching, is an emerging conceptualization of what good college teaching entails, focusing on the complex interactions between subject matter, teacher, learner, and context. Neumann (2014) claims that cognitively responsive teaching: (1) engages students deeply and intentionally in core subject matter ideas embedded in the discipline (Neumann 2005, 2014; Shulman 2004a, b); (2) connects students' learning of new subject matter ideas to their prior knowledge, both personal and cultural (Bransford et al. 2000; Ladson-Billings 1995); and (3) supports students whose prior knowledge clashes or challenges their learning of the new material (Bransford et al. 2000; Gonzalez et al. 2005).

Neumann's (2014) first claim suggests that good teaching facilitates students' encountering of subject matter ideas. Neumann argues that subject matter matters (Bransford et al. 2000; Dewey 1902, 1916; Shulman 2004a, b), and as such, student learning requires interaction with subject matter ideas culled from a discipline. This claim is based considerably on Shulman (2004a, b) theory of pedagogical content knowledge that contends that good teaching moves beyond having expertise in a content area—and also that good teaching moves beyond having expertise in broad-based pedagogical practices. Instead, Shulman (2004a, b) suggests that good teaching entails understanding that certain pedagogy works better with certain content as culled from the discipline. Pedagogical content knowledge supports the idea that expert teachers understand how students react to course content taught in certain ways, and therefore use this insight to make decisions about sequencing content, the use of different pedagogical practices, and the unfolding of content ideas across the class session and the course.

Moving forward, Neumann's (2014) second claim proposes that good teaching requires connecting students' learning of new course ideas with their prior knowledge. Rooted in previous research that suggests that students' learning of new ideas surfaces prior knowledge (Bransford et al. 2000; Dewey 1902; Shulman 2004a, b), this claim suggests that good teachers know how to engage students in conversations that bring forth their prior knowledge (Neumann 2014). Prior knowledge encompasses more than what the students know about the course content prior to the course (such as a pre-test), but it also includes the lived experiences and understandings, often cultural, that shape the students' ways of framing the course material. For example, in learning about the history course of the cold war, a student from Russian heritage may have prior knowledge that the professor would need to uncover, in depth, in order to proceed with learning.

Lastly, Neumann's (2014) third claim suggests good teaching involves supporting students' changing views. In this claim, good teachers provide opportunities for students to express being challenged by the juxtaposition of old and new ideas (Neumann 2014). This claim is rooted in the notion that good learning happens when a student reconciles differences between prior views and beliefs and new subject matter ideas from the course (Bransford et al. 2000; Shulman 2004a, b). While students' prior knowledge in some cases may aid learning, in others, their prior knowledge can serve as a barrier. Returning to the example of the history course on the cold war, the lived history of the cold war may have been very different for a student whose family is from Russia than a student whose family was living in the U.S. during the war. A professor can support both students in changing views by providing space for any dissonance between old views and new course ideas, and

then supporting students both cognitively and emotionally through the dissonance that arises (Neumann 2014).

## Methods

This study uses data from a multi-institutional quantitative observational study of 587 college classrooms. By “quantitative observation” we refer to an observation protocol that uses a closed-ended, highly structured rubric and coding scheme with raters who have been specifically trained to rate according to the conceptual framework used in this study (Johnson and Christensen 2010; Stallings and Mohlman 1988; Waxman et al. 2004). A benefit of using this method is that external raters, with knowledge of teaching and learning in higher education, can appraise the use of instructional practices in the classroom rather than relying on self-reported practices by instructors or students. Additionally, as Hora and Ferrare (2014) describe, this systematic observational method allows for raters to witness the subtle dynamics among students, faculty, course content, and context by being present in the classroom.

## Sampling

We purposively selected nine institutions to examine course practices across varying institutional types (2 research, 2 comprehensive, 5 liberal arts) and levels of selectivity (highly selective through broad access). Although these institutions are not generalizable to all undergraduate institutions, they represent a broad swath of institutions. To assist in understanding the transferability of results, we present the characteristics for these institutions in Table 1.

Within institutions, we employed random sampling of undergraduate courses, stratified by faculty category, class size, and discipline. Prior higher education literature has linked these factors to teaching quality or learning outcomes (Astin 1993; Johnstone and Maloney 1998; Pascarella and Terenzini 2005; Toth and Montagna 2002). We also weighted our sample of courses by seats within class size, ensuring that very small classes did not have an undue influence on our findings. While most classes were small, enrollment was clustered in larger classes, so a sample that considered the number of seats in each course better represented what a majority of students would experience at these institutions. We selected 350 courses from each institution, and faculty had the option to consent to include each course that was selected. Thirty-four percent of faculty who were invited to participate for one or more courses agreed to participate (similar response rate to national faculty survey rates). We were not able to observe all consented courses based on scheduling and observer constraints—among consented courses, we purposively selected courses that contributed to the representativeness of our sample along our strata. Observed courses were representative of course mode (online vs. on site), but slightly over-represented by tenure line faculty. As expected based on our sampling scheme, larger courses were over-represented due to weighting by enrollment. We present the representation of observed courses to the population of courses at the nine institutions in Table 2.



**Table 1** Characteristics of institutions

Variable	Institution description										
	High prestige public university	Top 100 national universities	High prestige private university	High prestige national liberal arts college	Top 25 national liberal arts college	Small liberal arts college	Urban liberal arts university	Rural liberal arts college	Suburban liberal arts college	Public state university	Rural multi campus regional university
Ranking and category	Top 50 national universities	Top 100 national universities	Top 100 national universities	Top 25 national liberal arts college	Top 25 national liberal arts college	50–100 national liberal arts college	Top 50 regional universities	>100 national liberal arts college	>100 national liberal arts college	Unranked regional universities	Unranked regional colleges
Carnegie classification	RU/VH	RU/VH	RU/VH	Bac/A&S	Bac/A&S	Bac/A&S	Master's	Bac/A&S	Bac/A&S	Master's	Bac/Assoc.
Size	Large	Medium	Medium	Small	Small	Small	Small	Small	Small	Medium	Medium
Small < 3000; medium < 10,000; large > 10,000											
SAT	High	Medium	Medium	High	High	Medium	Medium	Low	Low	Low	Low
High > 1250; medium = 1100–1250; low < 1100											
Graduation rate	High	High	High	High	High	Medium	Medium	Low	Medium	Low	Low
High > 80 %; medium > 65 %; low < 65 %											
Tuition	High	High	High	High	High	High	High	High	High	Low	Low
High > \$30,000; medium \$20,000–\$30,000; low < \$30,000											

**Table 2** Representativeness of observed courses compared with the population of courses at selected institutions

Characteristics	Observed courses		All courses	
	N	%	N	%
Class size				
5–25	335	57.1	4596	73.6
26–75	208	34.4	1450	23.2
>75	44	7.5	200	3.2
Tenure line <sup>a</sup>				
NTT	132	27.7	1690	40.1
TT	345	72.3	2525	59.9
Mode				
On site	562	95.7	5924	95.3
Online	25	4.3	295	4.7

<sup>a</sup> Two institutions did not report tenure status

*Note.* Classes in the sample were stratified by size and then weighted by seat to allow for larger courses to be selected at a greater rate than smaller classes

## Procedures

Site teams of seven to ten extensively trained observers visited each institution for 1 week during mid-semester. Observers participated in an observation training of approximately 30 hours. Observers were required to pass a certification process to demonstrate their knowledge of the conceptual frameworks and the observation procedures. They also were required to pass an inter-rater test to demonstrate reliability with ratings from the principal investigator. For each course, two observers rated the entire duration of one class period (with some exceptions due to scheduling conflicts where one rater observed).

## Data Sources

We created rubrics for the class observations to measure active learning, passive learning, and the three facets of Neumann's (2014) cognitively responsive teaching framework. The development of the rubrics was assisted by both content experts (in teaching quality) as well as methodological experts (in survey and rubric design), who tested the rubrics for content and response process validity. Active learning and cognitively responsive teaching were operationalized for this study as the presence or absence of seven course behaviors that mapped onto the theoretical framework. Specifically, we considered three broad categories of active learning techniques: class activities, class discussion, and student questions. Class discussion was observed in a class session as a free-flowing back and forth conversation among students or between students and the professor. We distinguished this from class activities that were more structured—for example, a faculty member could have given instructions for students to complete a task or set of related tasks during the class. A final form of active learning we included was when students asked questions of the professor during the class. Arguably, this is the least “active” technique, but still indicated student engagement and some degree of interactivity during the class session. In addition to the three active learning practices, we included lecture in our model. The lecture was observed in a class session as the instructor presenting course material to the students without interaction among students or between the instructor and students. This way to define lecture is similar to the way it is defined in the Freeman et al. (2014) study of active learning and lecture practices as “continuous exposition by the teacher.” Finally, we

included three teaching practices that mirror the three claims of good teaching in Neumann's framework on cognitive responsiveness: the orchestration of in-depth subject-matter ideas, understanding students' prior knowledge, and supporting students by bridging their prior knowledge to the new core ideas of the course. In order to demonstrate how the rubric on Neumann's (2014) framework of cognitively responsive teaching might manifest in specific course contexts, we provide an example of rubric tuning to a fictitious course in research methods in Table 3.

We present the seven behaviors, their operationalization, their mapping onto the theoretical frameworks, and their descriptive statistics in Table 4. We investigated patterns of courses across these seven dichotomous behaviors.

## Inter-rater Reliability

To calculate the inter-rater reliability of the observation data, we used a one-way, absolute, average-measure, mixed-effects intra-class correlation (ICC) calculation, which is appropriate for the current study given: the ordinal nature of the observation categories, the classes were random but the coders were fixed (i.e. we are not generalizing to a larger pool of raters outside the study), we used more than one rater per class, and we had an interest in the consistency of the absolute value of the ratings (Hallgren 2012). For observation, the ICC across all items was 0.705, with ICCs of sub-scales ranging from 0.664 to 0.787 (Cicchetti's 1994 cut off values: 0.60–0.74 good;  $\geq 0.75$  excellent).

## Disciplinary Considerations

Given that we studied college teaching in the context of disciplinary core ideas, we placed observers, as much as possible, in courses that matched their disciplinary backgrounds. As such, observers may be able to understand core concepts and how the frameworks apply within each discipline. We also did post hoc analyses to determine the extent to which disciplinary expertise matters in the ratings. We investigated whether the inter-rater reliability would be greater in courses where both observers had subject matter expertise compared with observers where only one had substantive subject matter expertise. We found that although the pairings of observers who both had expertise had slightly higher inter-rater reliability compared to pairing one expert observer with one non-expert observer (ICC = 0.737, 0.715, respectively), both pairings had good inter-rater reliability.

## Analyses

While the literature on teaching practices has adopted a variable-centered approach, we opted for latent class analysis (LCA) instead. Variable centered approaches such as factor analysis assume that the relation among variables is the by-product of a homogenous population. The end result is scales, quantifying the intensity of the variable. This homogeneity condition is not tenable in dealing with discrete or categorical events as the presence or absence of teaching behaviors in particular course setting (e.g., Masyn and Nylund-Gibson 2012; Wang and Wang 2012). Instead, LCA assumes that the heterogeneity of observed patterns of behaviors may be the product of subpopulations or latent

**Table 3** Cognitively responsive teaching tuning example, fictitious course in quantitative research methods in psychology

	Very ineffective example	Very effective example
Subject matter knowledge	<ul style="list-style-type: none"> <li>• The instructor introduces the definition of research, but does not situate it more broadly in the discipline of psychology</li> <li>• The instructor does not demonstrate to the class how the core ideas of today's class are inter-related with other classes in the semester</li> <li>• The instructor includes only one perspective on the core ideas</li> <li>• The instructor does not let students explore the limits and constraints of the core ideas for the day's class</li> <li>• For example, the instructor defines the concept of objectivity without letting students explore the assumptions in the field that undergird this concept</li> </ul>	<ul style="list-style-type: none"> <li>• The instructor defines, in depth, core ideas in quantitative research and includes ideas about the assumptions to quantitative methods from the discipline of psychology</li> <li>• The instructor explains these core ideas in multiple ways and using multiple pedagogies</li> <li>• The instructor maps out how the course and core ideas fit into psychology as a whole and the interconnections among sub-fields</li> <li>• The instructor discusses the progression of the course and how the ideas of this class are interconnected and build off other classes during the semester</li> </ul>
Prior knowledge	<ul style="list-style-type: none"> <li>• The instructor assumes that the core ideas can be learned similarly by all students</li> <li>• The instructor uses an example of quantitative research that is not relevant to students' lives</li> <li>• The instructor presents the core ideas, such as "objectivity," without seeking to understand what students know about this core idea and related concepts from their own lives</li> <li>• The instructor does not give students an opportunity to consider how their own experiences shape their understanding of research</li> </ul>	<ul style="list-style-type: none"> <li>• Instructor surfaces and explores students' prior knowledge</li> <li>• For example, the instructor brings forth examples that explore the assumptions of what quantitative research is in psychology</li> <li>• Instructor asks students to relate their own epistemological understandings to those of the researchers in the examples</li> <li>• The instructor might ask students how they came to have these understandings of the meaning of "research" based in their own lives</li> </ul>
Supporting changing views	<ul style="list-style-type: none"> <li>• The instructor does not explore whether students' prior understandings of research conflict with the new core ideas of quantitative methods</li> <li>• The instructor does not give students space to explore the dissonance they may feel about the conflicting views in their own lives and the course</li> <li>• The instructor does not support the emotional or intrapersonal features of this dissonance, for example in their math identity—"I am not good at math so I can't do research")</li> <li>• The instructor does not draw connections (cognitively) between old conceptions from students' lives and new course ideas</li> </ul>	<ul style="list-style-type: none"> <li>• The instructor creates a discussion opens space for students to show how they are challenged by new ideas and any dissonance that may arise</li> <li>• The instructor notes that some students are grappling with fears of math that are blocking them from exploring a deeper understanding of quantitative methods</li> <li>• In response to this dissonance, the instructor pulls out part of a students' personal history that shaped their understanding of "math" and connects it with the new assumptions of quantitative methods, thereby supporting the student in moving towards a new understanding</li> </ul>

classes (Geiser 2012; Masyn 2013; Wang and Wang 2012). In other words LCA is a statistical method for finding subtypes of related cases (latent classes) according to their observed values on a set of categorical or nominal indicators in cross-sectional data (Geiser 2012; Masyn and Nylund-Gibson 2012). While cluster analysis has been traditionally used to discover subpopulations of individuals, it has several limitations. Above anything else

**Table 4** Operationalization and descriptive statistics for seven course behaviors

Construct	Operationalization	Example observed behaviors
Active learning	Class discussion (Y = 45.8 %/N = 54.2 %) Class activities (Y = 61.5 %/N = 38.5 %) Student questions (Y = 80.1 %/ N = 19.9 %)	<ul style="list-style-type: none"> <li>• Back and forth conversation between instructor and students or among students about the course content</li> <li>• A structured activity where students engaged with the course content (e.g., case studies, clickers, group work)</li> <li>• Students asked individual questions of the instructor about the course content</li> </ul>
Passive learning	Lecture (Y = 85.5 %/N = 14.5 %)	<ul style="list-style-type: none"> <li>• A presentation or recitation of course content by the faculty member to all students in the class</li> </ul>
Subject matter knowledge	At least “Somewhat Effective” = yes; Ineffective or Very Ineffective = no (Y = 85.5 %/N = 14.5 %)	<ul style="list-style-type: none"> <li>• The instructor introduced in depth one or more concepts that are central to the subject matter of the course</li> <li>• The instructor created multiple representations of the “core ideas”</li> <li>• The instructor introduced students to how ideas play out in the discipline/field</li> </ul>
Students’ prior knowledge	At least “Somewhat Effective” = yes; Ineffective or Very Ineffective = no (Y = 56.5 %/N = 43.5 %)	<ul style="list-style-type: none"> <li>• The instructor surfaced students’ prior knowledge about the subject matter “core ideas”</li> <li>• The instructor worked to understand students’ prior knowledge about the subject matter “core ideas”</li> </ul>
Supporting changing views	At least “Somewhat Effective” = yes; Ineffective or Very Ineffective = no (Y = 50.5 %/N = 49.5 %)	<ul style="list-style-type: none"> <li>• Providing a space for students to encounter dissonance between prior knowledge and new course material</li> <li>• Helping students to realize the difference, similarities and sometimes conflict, between prior knowledge and new subject matter ideas</li> </ul>

Note Neumann’s (2014) framework was rated on a scale of 1 “Very Ineffective” to 5 “Very Effective;” “Somewhat Effective” was the midpoint (3) on the scale

cluster analysis is not an inferential technique (Wang and Wang 2012): the number of clusters is determined by rules of thumb and examination of graphs, which interject a great level of subjectivity on the part of the analyst. On the other hand, LCA is an inferential method allowing the analyst to compare alternative models of latent classes (Geiser 2012; Masyn 2013; Wang and Wang 2012).

Latent class analysis requires meeting three conditions: having a relatively small number of behaviors, having a sufficient sample size, and having a missing data pattern completely at random-MCAR (Masyn and Nylund-Gibson 2012). The sample of 587 courses is sufficiently large to analyze classes resulting from a hypothetical set of 49 combinations associated with our seven behaviors. Moreover, the condition that the probability of missing data on each of the 49 combinations as occurring completely at random was also met ( $\chi^2_{251} = 93.5, p = 1.0$ ). We ran our LCA analyses in MPlus statistical software (version 7.3).

Choosing the number of latent classes of courses that represent the behaviors under examination follows a step-wise process that combines statistical model-fit indicators with

model usefulness indicators based on classification quality and theoretical underpinnings related to the substantive interpretability of the classes (Masyn 2013; Wang and Wang 2012). We started with a single class model, testing the hypothesis of homogeneity of teaching classes, and then increased the number of classes by one each time to examine an alternative hypothesis of heterogeneity of teaching practices. To determine model fit, we used a combination of indicators as recommended by the LCA literature. Those included: the likelihood ratio Chi squared ( $\chi^2_{LR}$ ) statistics, BIC, BIC adjusted, AIC, and the adjusted Lo–Mendell–Rubin ( $\chi^2_{LMR-LRT}$ ) statistics.<sup>1</sup> To determine model usefulness, we examined the entropy ( $E$ ) statistics and the classification quality as determined by the average latent class probability for the most likely latent class membership by latent class. We also studied conditional probabilities of behaviors in each class to determine whether the classes were distinct and homogeneous in terms of course behaviors. Besides statistical indicators, we brought interpretability considerations to aid in the selection of the final model as suggested in the literature (e.g., Geiser 2012; Wang and Wang 2012).

## Results

Table 5 reports the results of six alternative cluster models of seven teaching behaviors. As we added classes, one at a time to the model, we examined the absolute and relative fit indices and considered whether the model significantly increased the fit as evidenced by the adjusted differences in  $\chi^2_{LMR-LRT}$  (see last column in Table 5). We also brought to bear our knowledge of the literature in guiding our model selection.

The 1-class model, corresponding to the hypothesis that there is no heterogeneity in teaching behaviors across the 587 courses, was rejected. The single class model has the largest maximum likelihood (−2364), BIC (4772), BIC-adjusted (4749) and AIC (4741) among all six models tested, suggesting that this model fits the data worse than all other models. In other words, there is heterogeneity in the way courses cluster together in our sample. While the two-class model and three-class model were improvements upon the one-class model, as indicated by lower BIC, BIC-adjusted and AIC, they did not fit the data as well as the four and five class models. We examined the fit of the four and five class models closely, given their similar fit values (see Table 5). At close inspection, we found support for a latent class model consisting of five classes. The likelihood ratio statistics for the five-class model is non-significant ( $\chi^2_{LR} = 96.9, p > 0.05$ ), suggesting it adequately fits our data. The five-class model also has the lowest AIC value (4241), and its entropy value of 0.821 is above the 0.80 threshold meeting the condition of being a reliable model (Wang and Wang 2012). While the four-class model has lower BIC values and a higher entropy value in relation to the five-class model, the BIC-adjusted value for the five-class model represents only a slight worsening of fit relative to the four-class model. Moreover, the adjusted  $\chi^2_{LMR-LRT}$  value test, comparing the four-class model versus the five-class model (see last column in Table 5), was statistically significant, indicating that adding the fifth class indeed increases the model fit. In addition to meeting the fit criterion, each component of the five-class model has a high degree of internal consistency as shown in

<sup>1</sup> Asparouhov and Muthén (2012) warn against relying on simple differences of  $\chi^2_{LR}$  to examine incremental fit in testing a  $k-1$  class versus a  $k$  class model. Instead, they recommend using the  $\chi^2_{LMR-LRT}$  statistics, which is estimated using the bootstrap procedures contained in Mplus TECH14 option.

**Table 5** Fit and modification fit indices for alternative cluster models of seven teaching behaviors

Model	LL	npar	$\chi^2_{LR}$ (df) (p value)	E	BIC	BIC-adjusted	AIC	Adj $\chi^2_{LMR-LRT}$ Ho: <i>k</i> classes versus H1 <i>k</i> + 1 classes (df) ( <i>p</i> -value)
1-Class	-2363.6	7	617.7 (119) ( <i>p</i> < 0.001)	-	4771.7	4749.5	4741.1	-
2-Classes	-2130.7	15	194.2 (111) ( <i>p</i> < 0.001)	0.777	4357.1	4309.5	4291.5	<b>465.6 (8)</b> ( <i>p</i> < 0.001)
3-Classes	-2103.0	23	138.2 (103) ( <i>p</i> < 0.05)	0.865	<b>4352.7</b>	<b>4279.7</b>	4252.1	<b>55.4 (8)</b> ( <i>p</i> < 0.01)
4-Classes	-2092.2	31	<b>118.2 (96)</b> ( <i>p</i> = 0.06)	<b>0.899</b>	4381.9	4283.5	4246.3	<b>21.8 (8)</b> ( <i>p</i> < 0.05)
5-Classes	-2081.6	39	<b>96.9 (88)</b> ( <i>p</i> = 0.24)	0.821	4411.8	4287.9	<b>4241.1</b>	<b>21.2 (8)</b> ( <i>p</i> < 0.05)
6-Classes	-2075.8	47	<b>85.4 (80)</b> ( <i>p</i> = 0.32)	0.800	4451.2	4302.0	4245.6	11.6 (8) ( <i>p</i> = 0.60)

Note Item probabilities at or greater than 0.65 are bolded to indicate a high degree of class homogeneity

**Table 6** Average latent class probabilities for the most likely latent class membership (row) by latent class (column)

	Comprehensive	Traditional Lecture	Active Lecture	Integrative Discussion	Active Only
Comprehensive	0.940	0.007	0.049	0.002	0.002
Traditional Lecture	0.054	0.736	0.210	0.000	0.000
Active Lecture	0.047	0.103	0.850	0.000	0.000
Integrative Discussion	0.124	0.000	0.000	0.871	0.005
Active Only	0.027	0.000	0.000	0.000	0.973

*Note* The rows report the conditional probability of class membership for each of the seven behaviors within each of the latent class

Table 6. Finally, the classes identified by the five-class model, as shown below, are interpretable. Collectively, this evidence lends strong support for the five-class model.

Table 7 presents the results of the LCA, revealing five distinct clusters. The rows report the conditional probability of class membership for each of the seven behaviors within each of the latent classes. Following Masyn's (2013) recommendation, we judged behavior membership to a particular class based on 0.70 and 0.30 thresholds (we also highlight probabilities that are close to these cut-offs).

**Table 7** Unconditional and conditional probabilities for the 5-class LCA model ( $N = 587$ )

	Latent class				
	Comprehensive ( $N = 323$ )	Traditional Lecture ( $N = 113$ )	Active Lecture ( $N = 112$ )	Integrative Discussion ( $N = 10$ )	Active Only ( $N = 29$ )
	Unconditional probability for each class				
	0.550	0.193	0.191	0.017	0.049
	Conditional probability of membership of each behavior within each latent class				
1. Class activities	<b>0.706</b>	<i>0.147</i>	<b>0.702</b>	0.571	<b>0.791</b>
2. Class discussion	<b>0.650</b>	<i>0.006</i>	0.371	<b>0.825</b>	<i>0.152</i>
3. Student questions	<b>0.932</b>	0.583	<b>0.802</b>	<i>0.000</i>	0.351
4. Lecture	<b>0.853</b>	<b>1.000</b>	<b>1.000</b>	<i>0.000</i>	<i>0.000</i>
5. Subject-matter knowledge	<b>0.982</b>	<b>0.795</b>	<b>0.735</b>	<b>0.690</b>	<i>0.253</i>
6. Students' prior-knowledge	<b>0.912</b>	<i>0.222</i>	<i>0.111</i>	0.398	<i>0.103</i>
7. Supporting changing views	<b>0.865</b>	<i>0.000</i>	<i>0.094</i>	<b>1.000</b>	<i>0.000</i>

*Note* Item probabilities at or greater than 0.65 are bolded and those less than 0.3 are italicized to indicate a high degree of class homogeneity



Class 1, labeled ‘Comprehensive’, represents a category of courses that was likely to enact a variety of teaching practices, including all seven practices investigated in this study. This class was the largest among the five classes representing 55 % of the sample. We labeled this class as ‘Comprehensive’ because the courses were likely to use teaching practices aligned with both frameworks we examined in this study (active learning and cognitively responsive teaching), as well as lecture. Notably, contrary to the literature that pins active learning as opposed to lecture (Freeman et al. 2014), this latent class enacted both lecture *and* active strategies. Additionally, instructors of the ‘Comprehensive’ courses delved deeply into subject-matter knowledge and students’ prior knowledge.

Class 2, labeled ‘Traditional Lecture,’ was most characterized by lecturing and subject matter knowledge. This class represented 19.3 % of the courses in our sample. In particular, these courses used traditional and teacher-centered modes of instruction, and did not enact active learning or cognitively responsive teaching practices. The internal consistency of this class, while still in the acceptable range, was lower than other classes (0.736) and had some overlap with class 3, ‘Active Lecture’ (Table 6). Given the overlap among the ‘Traditional Lecture’ and the ‘Active Lecture’ classes, there may have been certain courses that were emerging from traditional passive pedagogies and moving in the direction of becoming more active in their teaching practices.

Class 3, labeled ‘Active Lecture’, included a combination of lecture and active learning practices. Representing 19.1 % of courses, this class enacted traditional lecture and also incorporated class activities and student questions. Like the ‘Traditional Lecture’ group of courses, the ‘Active Lecture’ courses demonstrated teaching with in-depth subject-matter knowledge. However, these courses were not characterized by teaching practices that delved into students’ prior knowledge, according to Neumann’s (2014) framework.

Class 4, labeled ‘Integrative Discussion’, is the smallest among the five classes, accounting for only about 2 % of the courses. This class was selective in its practices. For example, ‘Integrative Discussion’ courses enacted active learning by using class discussion, but not activities or student questions. These courses also focused in on subject-matter knowledge and supporting students in making connections between their prior knowledge and the new course ideas. In this way, these courses may have used discussion to help students integrate their prior knowledge with new course ideas.

Class 5 represented only about 5 % of the courses examined. We called this class ‘Active Only’ because it enacted active learning by getting students engaged in class activities, but it was not organized around either in-depth subject matter expertise or the students’ prior knowledge. Courses in this class were unlikely to include class discussions, lecture, student questions, or any of Neumann’s (2014) three cognitively responsive teaching practices (subject matter knowledge, students’ prior knowledge, or supporting changing views). Out of the seven teaching practices, this class was only characterized by class activities, leaving us to question the focus and direction of these activities.

We closely examined the evidence for the final two latent classes given that they included only a small proportion of the courses in this study. While they are small, the ‘Integrative Discussion’ and ‘Active Only’ latent classes had very high internal consistency (0.871, 0.973 respectively). Additionally, these two latent classes of courses were consistent in both the 4-class and 5-class models that we tested, lending support for them as distinct latent classes. Finally, these classes represented meaningful categories of courses given the theory on active learning and cognitively responsive teaching, which we consider more fully in the discussion section. While we believe that the evidence from the model supports the inclusion of these smaller latent classes, future studies should attempt to replicate these results with larger samples.

## Limitations

This study has certain noteworthy limitations. This study investigates teaching practices across disciplinary contexts. The classes of courses may change if the study was conducted within one particular discipline, and we see this as an important area of further research. However, given that this is an initial examination of a broad number of college classrooms, this study provides a view across classes of courses and the practices they enact. The way this study considers college course practices across disciplinary contexts is similar to other multi-institutional studies of college teaching (e.g., Carini et al. 2006; Umbach and Wawrzynski 2005). Secondly, the study is of 587 courses, but we only visited one class session, mid-semester. This is a limitation given that teaching unfolds across a semester and the ordering of course ideas is germane to pedagogical content knowledge (Major and Palmer 2006; Shulman 2004b). Yet, the study does reveal a snapshot in time of the practices across nine institutions—and perhaps, in the aggregate, this gives a sense of the “mid-semester” practices taking place on these institutions, even if individual course practices may vary from course to course. Another possible limitation is the reactivity involved in observation (participants may react differently while observed; Anderson and Burns 1989; Jacob et al. 1987). We attempted to limit these biases by instructing observers to be as unobtrusive as possible, reassuring participants about the confidentiality of the data. Finally, this study includes a sample of only nine institutions and is not necessarily generalizable to the broader population of higher education institutions. We provide information about the institutions in our sample in Table 1 to assist in interpreting whether the results transfer to other higher education settings. Given these reasons, we consider our study to be an exploratory examination of the complexity of teaching practices that occurs among courses.

## Discussion

This study used latent class analysis (LCA) to determine whether there were patterns of courses mid-semester based on teaching behaviors from active learning theory and cognitively responsive teaching theory. In prior literature, studies examined which courses enacted individual practices, such as active learning techniques (Cabrera et al. 2001, Freeman et al. 2014; Hora and Ferrare 2014). However, by using LCA, we were able to learn whether groups of courses enacted sets of teaching practices (e.g., active learning in combination with lecture and cognitive responsiveness). Our study found that five groups of courses enacted specific patterns of educational practices, and with strong consistency within each group. In essence, this means that the courses within the study are heterogeneous in that there are five groups of courses, but within each group, these courses act similarly. These results characterize a strong LCA model (Wang and Wang 2012).

We see two main contributions of this study to the literature on college teaching and learning. First, understanding groups of courses and the teaching practices that characterize them gives insight into how teachers are, or are not, enacting both traditional and newer college teaching frameworks, and yields a multidimensional perspective on teaching practices that a variable centric approach cannot depict (Wang and Wang 2012). Second, surprisingly few multi-institutional studies focus at the course level when examining college teaching. Multi-institutional studies of college teaching have mainly surveyed students about their overall course experiences (Kuh 2001; Reason et al. 2006) or faculty

about their overall teaching practices (DeAngelo et al. 2009; Feldman 1989a, b; Umbach and Wawrzynski 2005). This study observed almost 600 college classes to view the educational practices as they unfold within courses.

Returning to the theoretical contributions of this study, within the literature base on college teaching, we see a continued evolution of understanding effective practices. Although the literature base is broad and non-linear, we see a general movement from didactic teacher-centered paradigms (i.e., lecture) to active learning and student-centered paradigms to responsive paradigms (i.e. cognitively, culturally, or context responsive). We see artifacts of each of these movements in the classes of courses found our study. We also find that the classes of courses in this study depart from the traditional ways of describing these frameworks by uncovering the complex interplays across these frameworks in the way they manifest in the classroom.

The literature on college teaching has largely critiqued the traditional paradigm focusing on the instructor as the purveyors of expertise to students (as receivers) through a recitation or lecture (McKeachie and Kulik 1975). Our study revealed that one class of courses, the ‘Traditional Lecture’ class, still enacted these more traditional teaching practices, based on a teacher and expertise-centered paradigm. These courses, at least in the one class session observed, appear to be unsuccessful in active learning techniques as they did not engage students in class activities, class discussion, or student questions. These courses also did not appear to take into account the cognition of the students by incorporating teaching behaviors using students’ prior knowledge or connecting that prior knowledge to the core ideas of the course. It is possible that these courses are “stuck” in older paradigms of teaching; they have not taken up the call to active learning and engagement, nor have they delved into the minds of students through Neumann’s (2014) framework. Another possibility given that we observed one class mid-semester is that these courses used this particular class section for lecture, and may enact active learning and cognitively responsive teaching in subsequent class sessions. Future research that investigates these ‘Traditional Lecture’ courses across the semester would provide additional insight into how these practices unfold over time.

The Traditional Lecture class represented about 19 % of the courses in this study. In this respect, our findings run contrary to earlier studies that found that lecture was still a predominant teaching strategy in college teaching even after the considerable evidence of active learning practices took hold (Freeman et al. 2014; Hora and Ferrare 2013; Koljatic and Kuh 2001; Lammers and Murphy 2002). This discrepancy between our findings and the literature base can be traced back to the different analytical methods used in examining teaching practices. If a variable-centric rather than latent class approach had been used, we would have reported that 85.5 % of courses in our study had included a lecture, a finding largely consistent with prior literature using variable-centered methods for studying teaching behaviors. By using LCA, we were able to detect patterns of class practices that certain courses enact. The ‘Traditional Lecture’ group was the only class of courses to enact lecture with subject-matter expertise, but no other active learning or cognitively responsive practices.

Moving beyond traditional lecture, for decades, evidence accumulated that involving students in their education through active learning and student-centered practices was a more effective college teaching practice and contributed significantly to student learning gains (Cabrerera et al. 2001; Colbeck et al. 2001) when compared to traditional lecture and teacher-centered approaches. More recently, Freeman et al. (2014) did a meta-analysis of college teaching studies and concluded that lecture was a “pariah” in college teaching, and should be left behind in favor of “active” practices. Our study reveals good news, in that

four of the five classes of courses in our study were likely to include active learning techniques. However, the LCA analysis moves beyond this approach by examining the sets of practices that are enacted together with the active learning practices. This approach, perhaps, provides a more dynamic picture of the college classroom than the one painted by variable-centered methods.

For example, the ‘Active Lecture’ group uses lecture, student questions, and class activities—so is ‘active’ by engaging students in their learning process—but also utilizes the traditional lecture approach. Yet, this group of courses does not surface students’ prior knowledge or connect that knowledge to the new core ideas of the course. In essence the ‘Active Lecture’ courses integrated the active learning framework into their repertoire, but do not yet enact cognitively responsive teaching. Similarly, the ‘Active Only’ class of courses was likely to enact class activities, but unlikely to portray in-depth subject matter ideas or respond to students’ cognitions of the core ideas. We see these classes as potential evidence that the emphasis on singular categories (such as active learning/lecturing; Hora and Ferrare 2014) may be causing interactive education to take place without the associated meaning-making. Additionally, research using a variable centric approach cannot see heterogeneous groups of courses utilizing sets of practices. It, therefore, obscures which active learning courses or lecture courses also enact associated meaning-making using cognitively responsive paradigms.

Perhaps the most promising finding from our study was the ‘Comprehensive’ class, representing the largest proportion of courses in our study (55 %). The courses in the ‘Comprehensive’ class enacted active learning through class activities, class discussion, and student questions. While the findings of this study support active learning in the classroom, they also move beyond these binary categories. Contrary to prior literature that casts active learning strategies as favorable and lecture as problematic (Freeman et al. 2014), we find that the ‘Comprehensive’ courses both use lecturing and active learning practices. By using several modes of delivery (lecture, activity, discussion), these courses also are able to enact Neumann’s cognitively responsive framework. Out of all five classes of courses, the ‘Comprehensive’ courses, that use both active learning and lecture practices, were the only ones that also enacted all three of Neumann’s (2014) framework—in-depth subject matter expertise, understanding students’ prior knowledge, and bridging students’ prior knowledge with new core ideas of the course. Neumann (2014) contends, based on a framework from the learning sciences, that these three elements are necessary for learning. Lecturing may be an important ingredient of pedagogical content knowledge and connecting core ideas to students’ own understandings, under certain specific course contexts.

The final class in our analysis, “Integrative Discussion,” represents only 2 % of the courses in our sample. Yet, we see this class of courses as important in understanding how Neumann’s framework may be enacted in courses. These courses held class discussions and the instructor described in-depth subject matter ideas and helped bridge students’ prior knowledge to these core ideas. It may be that this particular class session mid-semester was devoted to discussing, in-depth, the integration of the course ideas to students’ prior knowledge. Perhaps class sessions earlier in the semester within the same course may have used other pedagogical techniques or focused on understanding the students’ prior knowledge of the core ideas. Perhaps the process of integrating core subject matter ideas with students’ prior knowledge requires in-depth discussion among students and instructor (Neumann 2014).

Our findings seem to support Hora and Ferrare’s (2014) conclusion regarding the policy approach to teaching: “In attempting to categorize teaching into simple dichotomous

groups (e.g., lecturing vs. interactivity), policy makers are provided with an inaccurate and coarsely grained perspective of teaching that does not reflect the realities of classroom practice” (p. 36). By contrast, our LCA analysis was able to uncover how teaching practices (lecture, active learning, cognitive responsiveness) work together or separately in classes of courses. Simply ‘becoming active’ without the associated meaning-making or framing from students’ lived experiences or the core ideas of the course may not be enough to produce learning. Likewise, there may be moments in courses where focusing in on certain kinds of modes or pedagogical approaches, such as seen in the ‘Integrative Discussion’ class of courses, may be warranted. In this way the duality of a theory that juxtaposes active against passive learning does not seem to fit the data in this study, but this warrants future research. While this was an important finding given that the literature base often draws upon this duality, we wonder whether more nuanced understandings of what “active learning” might mean could be more meaningful in terms of representing what college teachers practice in their courses. For example, we wonder if the rubrics had more specifically examined studio-based or problem-based learning (Kwan 2009; Little and Cardenas 2001) instead of a broader category of “class activities”—would this be more meaningful in the course clusters?

The findings of this study suggest further research into college classrooms to understand the complex dynamics that take shape between students, faculty, course content, and context (Campbell in press; Neumann 2014). There has been a great deal of attention on the High Impact practices touted by the American Association of Colleges and Universities (AAC&U), such as culminating experiences, service learning, and internships (Kuh 2008). In addition to these practices, which often focus on the co-curriculum or a specialized curriculum, our research points to investigating the complex phenomenon inside and across college classrooms. Additional studies could further research what the ‘Comprehensive’ courses look like in practice. Further research could consider the timing of these courses in the beginning, middle, and end of the semester and how these sets of practices change overtime. Future research could also investigate the necessary conditions for these class memberships—do certain institutional types or course characteristics, such as class size, discipline, or faculty characteristics breed ‘Comprehensive’ or ‘Active Only’ courses? Perhaps, most importantly, do these classes of courses predict certain learning outcomes? Based on the college teaching literature and broader literature from the learning sciences, the individual practices of active learning and understanding and integrating students’ prior knowledge have been found effective in student learning (Cabrera et al. 2001; Colbeck et al. 2001; Freeman et al. 2014; Neumann 2014). However, this study found that these practices are utilized together in groups of courses, which leads to the question of whether these groups predict outcomes. Finally, for the faculty who seem to be ‘stuck’ in the traditional lecture model, further research could investigate whether there are ways to enhance faculty understanding of these new frameworks and to help faculty make them actionable.

**Acknowledgments** This research was supported by a National Academy of Education/Spencer Foundation fellowship.

## References

Anderson, L. W., & Burns, R. B. (1989). *Research in classrooms: The study of teachers, teaching, and instruction*. Oxford: Pergamon.

- Arends, R., & Castle, S. (1991). *Learning to teach* (Vol. 142). New York: McGraw-Hill.
- Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago: University of Chicago Press.
- Asparouhov, T., & Muthén, B. (2012). Using Mplus TECH11 and TECH14 to test the number of latent classes. Mplus Web Notes: No. 14.
- Astin, A. W. (1993). *What matters in college? Four critical years revisited* (1st ed.). San Francisco: Jossey-Bass.
- Beach, A. L., Henderson, C., & Finkelstein, N. (2012). Facilitating change in undergraduate STEM education. *Change: The Magazine of Higher Learning*, 44(6), 52–59.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Braxton, J. M., Milem, J. F., & Sullivan, A. S. (2000). The influence of active learning on the college student departure process: Toward a revision of Tinto's theory. *The Journal of Higher Education*, 71(5), 569–590.
- Cabrera, A. F., Colbeck, C. L., & Terenzini, P. T. (2001). Developing performance indicators for assessing classroom teaching practices and student learning. *Research in Higher Education*, 42(3), 327–352.
- Campbell, C. M. (in press). An inside view: The utility of quantitative observation in understanding college educational experiences. *Journal of College Student Development*.
- Carey, K. (2014, July 28). Building a Better College Ranking System. Wait, Babson Beats Harvard? *The New York Times*.
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, 47(1), 1–32. doi:10.1007/s11162-005-8150-9.
- Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, 64, 723–733.
- Centra, J. A. (1973). Effectiveness of student feedback in modifying college instruction. *Journal of Educational Psychology*, 65(3), 395.
- Colbeck, C. L., Cabrera, A. F., & Terenzini, P. T. (2001). Learning professional confidence: Linking teaching practices, students' self-perceptions, and gender. *The Review of Higher Education*, 24(2), 173–191.
- DeAngelo, L., Hurtado, S., Kelly, K. R., Santos, J. L., & Korn, W. S. (2009). *The American college teacher: National norms for the 2007–2008 HERI faculty survey*. Los Angeles: Higher Education Research Institute, UCLA.
- Dewey, J. (1902). The child and the curriculum. In R. D. Archambault (Ed.), *On education* (pp. 339–358). Chicago, IL: University of Chicago Press.
- Dewey, J. (1916). The nature of subject matter. In R. D. Archambault (Ed.), *On education* (pp. 359–372). Chicago, IL: University of Chicago Press.
- Eagan, M. K., Stolzenberg, E. B., Berdan Lozano, J., Aragon, M. C., Suchard, M. R., & Hurtado, S. (2014). *Undergraduate teaching faculty: The 2013–2014 HERI Faculty Survey*. Los Angeles: Higher Education Research Institute, UCLA.
- Feldman, K. A. (1989a). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. *Research in Higher Education*, 30(2), 137–194.
- Feldman, K. A. (1989b). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30(6), 583–645.
- Feldman, K. A. (1997). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry & J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 368–395). New York: Agathon.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8410–8415.
- Geiser, C. (2012). *Data analysis with Mplus*. New York, NY: Guilford Press.
- Gonzalez, N., Moll, L. C., & Amanti, C. (2005). *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Guttenplan, D. (2014, June 1). Re-evaluating the college rankings game. *The New York Times*.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23.
- Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, 48(8), 952–984.

- Hora, M. T. (2014). Limitations in experimental design mean that the jury is still out on lecturing. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(30), E3024–E3024.
- Hora, M. T., & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of math and science undergraduate course planning and classroom teaching. *Journal of the Learning Sciences*, *22*(2), 212–257.
- Hora, M. T., & Ferrare, J. J. (2014). Remeasuring postsecondary teaching: How singular categories of instruction obscure the multiple dimensions of classroom practice. *Journal of College Science Teaching*, *43*(3), 36–41.
- Jacob, T., Tennenbaum, D. L., & Krahn, G. (1987). *Factors influencing the reliability and validity of observation data* (pp. 297–328). US: Springer.
- Johnson, B., & Christensen, L. (2010). *Educational research* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Johnstone, D. B., & Maloney, P. A. (1998). Enhancing the productivity of learning: Curricular implications. *New Directions for Higher Education*, *1998*(103), 23–34.
- Kaminer, A. (2013, October 27). Lists that rank colleges' value are on the rise. *The New York Times*.
- Koljatic, M., & Kuh, G. D. (2001). A longitudinal assessment of college student engagement in good practices in undergraduate education. *Higher Education*, *42*(3), 351–371.
- Kuh, G. D. (2001). Assessing what really matters to student learning inside the national survey of student engagement. *Change: The Magazine of Higher Learning*, *33*(3), 10–17.
- Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: American Association of Colleges and Universities.
- Kwan, A. (2009). Problem based learning. In M. Tight, H. K. Mok, J. Huisman, & C. C. Morphew (Eds.), *The Routledge international handbook of higher education*. New York: Routledge.
- Ladson-Billings, G. (1995). But that's just good teaching: The case for culturally relevant pedagogy. *Theory into Practice*, *34*(3), 159–165. doi:10.1080/00405849509543675.
- Laird, T. F. N., Chen, D., & Kuh, G. D. (2008). Classroom practices at institutions with higher than expected persistence rates: What student engagement data tell us. *New Directions for Teaching and Learning*, *2008*(115), 85–99.
- Lammers, W. J., & Murphy, J. J. (2002). A profile of teaching techniques used in the university classroom A descriptive profile of a US public university. *Active Learning in Higher Education*, *3*(1), 54–67.
- Lattuca, L. R., Voigt, L. J., & Fath, K. Q. (2004). Does interdisciplinarity promote learning? Theoretical support and researchable questions. *The Review of Higher Education*, *28*(1), 23–48.
- Little, P., & Cardenas, M. (2001). Use of “studio” methods in the introductory engineering design curriculum. *Journal of Engineering Education*, *90*(3), 309–318.
- Magner, D. K. (1997). Report says standards used to evaluate research should also be used for teaching and service. *Chronicle of Higher Education*, *44*(2), A18–A19.
- Major, C. H., & Palmer, B. (2006). Reshaping teaching and learning: The transformation of faculty pedagogical content knowledge. *Higher Education*, *51*, 619–647.
- Marsh, H. W., & Roche, L. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American educational research journal*, *30*(1), 217–251.
- Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. *The Oxford handbook of quantitative methods in psychology*, *2*, 551–611.
- Masyn, K. E. & Nylund-Gibson, K. (2012, August 15). *Introduction to LCA and finite mixture modeling using Mplus*. Johns Hopkins University, Bloomberg School of Public Health. Baltimore.
- McKeachie, W. J., & Kulik, J. (1975). Effective college teaching. *Review of Research in Education*, *3*, 165–209.
- National Survey of Student Engagement. (2015). *Engagement insights: Survey findings on the quality of undergraduate education—Annual results 2015*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- Neumann, A. (2005). To glimpse beauty and awaken meaning: Scholarly learning as aesthetic experience. *The Journal of Aesthetic Education*, *39*(2), 68–88.
- Neumann, A. (2014). Staking a claim on learning: What we should know about learning in higher education and why. *The Review of Higher Education*, *37*(2), 249–267.
- Neumann, A. & Campbell, C. M. (2016). Homing in on learning and teaching: Current approaches and future directions for higher education policy. In M. N. Bastedo, P. G. Altbach, & P. J. Gumpert (Eds.), *American Higher Education in the Twenty-First Century: Social, Political, and Economic Challenges* (4<sup>th</sup> ed.). Baltimore: Johns Hopkins University Press.

- Oleson, A., & Hora, M. T. (2014). Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices. *Higher Education*, 68(1), 29–45.
- Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students* (Vol. 2). In K. A. Feldman (Ed.). San Francisco, CA: Jossey-Bass.
- Reason, R. D., Terenzini, P. T., & Domingo, R. J. (2006). First things first: Developing academic competence in the first year of college. *Research in Higher Education*, 47(2), 149–175.
- Robertson, D. R. (2005). Generative paradox in learner-centered college teaching. *Innovative Higher Education*, 29(3), 181–194. doi:10.1007/s10755-005-1935-0.
- Rothkopf, E. Z. (1973). What are we trying to understand and improve? Educational research as *Leerlaufreaktion*. *Educational Psychologist*, 10(2), 58–66.
- Shulman, L. S. (2004a). *Teaching as community property: Essays on higher education*. In P. Hutchings (Ed.). San Francisco, CA: Jossey-Bass.
- Shulman, L. S. (2004b). *The wisdom of practice: Essays on learning, teaching, and learning to teach*. San Francisco: Jossey-Bass.
- Stallings, J. A., & Mohlman, G. G. (1988). Classroom observation techniques. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 469–474). Elmsford, NY: Pergamon Press.
- Toth, L. S., & Montagna, L. G. (2002). Class size and achievement in higher education: A summary of current research. *College Student Journal*, 36(2), 253.
- Umbach, P. D., & Wawrzynski, M. R. (2005). Faculty do matter: The role of college faculty in student learning and engagement. *Research in Higher Education*, 46(2), 153–184.
- Van Der Meij, H. (1994). Student questioning: A componential analysis. *Learning and Individual Differences*, 6(2), 137–161.
- Volkwein, J. F., Lattuca, L. R., Harper, B. J., & Domingo, R. J. (2007). Measuring the impact of professional accreditation on student experiences and learning outcomes. *Research in Higher Education*, 48(2), 251–282.
- Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Chichester, West Sussex: Wiley.
- Waxman, H. C., Tharp, R. G., & Hilberg, R. S. (Eds.). (2004). *Observational research in U. S. classrooms: New approaches for understanding cultural and linguistic diversity*. West Nyack, NY: Cambridge University Press.