Competing Institutional Logics And Teaching Effectiveness In Traditional And Online University Classrooms

Billy R. Brocato
Texas A&M University-San Antonio, billbrocato@tamu.edu

Oi-man Kwok
Texas A&M University, omkwok@tamu.edu

Follow this and additional works at: https://digitalcommons.kennesaw.edu/jpps

Recommended Citation
Available at: https://digitalcommons.kennesaw.edu/jpps/vol10/iss1/5
Competing Institutional Logics And Teaching Effectiveness In Traditional And Online University Classrooms

Cover Page Footnote
The authors wish to acknowledge the Texas A&M University Center for Integrating Research, Teaching and Learning (CIRTL) for providing a Teaching-as-Research Fellows Program Award that partially funded this research project. Sample data and statistical output is available upon request. The authors also would like to express their gratitude to the journal's anonymous reviewers for their insightful critique and helpful suggestions in an earlier draft of this manuscript.
INTRODUCTION

Relying on an institutional logics framework (Townley 1997), we offer a critical sociological perspective (Apple 2013) to the ongoing debate in academia over the effectiveness of student evaluations of instructors (SEI). To accomplish our task, we use a case study method to examine competing inter-institutional logics effecting U.S. postsecondary instructor ratings in traditional and online courses at a midsize public university’s college of humanities and sciences. This is an important because prior research has attributed SEI outcome differences to various instructor and student attitudinal, performative, and social factors without examining these outcomes in light of competing institutional conditions (Abrami, D'Apollonia, and Cohen 1990; Mentzer, Cryan, and Tellehaianot 2007; Kuzmanovic et al. 2014; Dodeen 2013; Young and Duncan 2014; Ryan 2015).

BACKGROUND

The effect of competing institutional logics—scholarly versus neoliberal—in U.S. post-secondary universities is playing out in a discursive space between the financialization and the cultural autonomy of the academy (Donoghue 2008; Hunsaker and Thomas 2014; Sladek 2014). As a meta-theory, institutional logics posit that actors rely on institutional rules to organize their required social interactions in formal settings; institutional logics are those processes that inform actors’ cognitions and behaviors in times of uncertainty (Thornton 2002; Dunn and Jones 2010). New institutional theorists have argued that the legitimacy of the isomorphic structure of institutions occurs as organizational actors coalesce around a specific set of practices that have proved to be essential to the organization’s sustainability and then is copied by other organizations. The isomorphic structure then is copied across organizational fields of interaction such as in public education institutions in the United States. Specifically, Wang (2016:349) has written that “Isomorphism describes the convergence of ‘organizational forms and practice’ in organizational fields. NI [new institutionalist] theorists argue that organizations tend to be alike because organizational actors unconsciously respond to the same set of institutionalizing forces in the social environment” [italics added for emphasis].

We believe actors’ ‘unconscious responses’ to institutional rules and practices is more about ‘maneuvering’ in a social setting by relying on an institutional logic that allows actors to arrive at sensible decisions aimed at maximizing their social position in times of certainty and uncertainty (see DiMaggio 1998:700-701). In this paper, we are concerned with delineating the site of competing institutional logics that structure actors’ behaviors in higher educations in a time of competing logics. Nowhere is this more obvious than in the decades-long shift to a neoliberal framework that has deemphasized the
public financing of education; especially undermining public universities’ attempts to maintain liberal arts programs and faculty autonomy (McCall 2000; Katsinas and Friedel 2010; Apple 2013).

The neoliberal agenda values free market competition through laissez-faire economic policies over state supported programs. Having gained credence in the United States during Ronald Reagan’s presidency and in Britain under Prime Minister Margaret Thatcher in the 1980s, neoliberal economic thought began its global ascendancy (defined later as globalization) through dominant international institutions such as the International Monetary Fund, the World Bank, and the World Trade Organization (Flew 2014).

Giroux (2014:17-18) has criticized neoliberalism’s penetration into the public sphere because it has aligned public institutions such as state colleges and universities with “the organizational trappings of medium-sized or large corporations” whereby “university presidents are now viewed as CEOs, faculty as entrepreneurs, and students as consumers”. The result has been a shift in traditional university administrators’ responsibilities away from their role as protectors of the academy’s cultural autonomy to public relations experts working to attract corporate sponsors as has become common in the for-profit post-secondary sector (Smith 2010; De Leo 2013).

More recently, Cottom (2017) has detailed how the commodification of higher education bolstered by a neoliberal economic structure has fostered the growth of technical schools and online universities’ ability to attract and serve marginalized or ‘high-risk’ student groups unable to meet the costlier tuitions of public colleges and more substantively, the lengthy time commitments needed to complete a bachelor’s degree. She has defined the for-profit sector as “Lower Ed” (italics in the original) in contrast to elite Higher Ed colleges (2017:12).

This rapid expansion and corporatizing of the Lower Ed sector has contributed to the meteoritic rise of dual-role colleges—scaled down brick-and-mortar facilities alongside expanded Web-based online degree programs. The most successful and recognizable Lower Ed entities are the University of Phoenix, DeVry University, Kaplan University, and Walden University (Breneman 2005; Cronin and Bachorz 2005; PBS Frontline 2010). In this new institutional sphere structured by a logic of neoliberalism, higher education has become a place where students are consumers, faculty are a cheap form of precariat labor, and each college’s academic department is perceived as a revenue-generating enterprise (Giroux 2014).

STUDENT EVALUATIONS OF INSTRUCTORS: AN INSTITUTIONAL DILEMMA

A major consequence of the corporatizing of American universities has been the coercive traction student evaluations of instructors (SEIs) have gained recently as public universities have begun to make public the results of
instructors’ assessments (Narayanan, Sawaya III, and Johnson 2014). This portends unapologetic valuations by a public that has little awareness of university teachers’ multi-purpose roles, the academy’s traditional cultural autonomy from economic and political persuasions, or of the differentiation of organizational roles and responsibilities within the academy’s professional classes (Ball 2008; Chin, Senter, and Spalter-Roth 2011).

The effect of SEIs on faculty tenure, the loss of intellectual freedom, and the re-imaging of students as clients thus exemplifies a growing conflict within American postsecondary institutions (Wiesenberg and Stacy 2008; Menand 2010; Schrecker 2010; Pană 2015; Saunders 2015; Bañuelos 2016). However, as both groups have struggled to find valid interpretations for well-documented disparities in SEI assessments (Miller and Chamberlin 2000; Riniolo, Johnson, and Sherman 2006; Kogan, Schoenfeld-Tacher and Hellyer 2010), we suggest that a key explanatory theory left understudied is the contextualizing effect of competing institutional logics on instructors and students decision making during evaluations (Bolliger and Martindale 2004; Nichols 2011; Myers and Claus 2012; Seddon 2014).

For example, Pană (2015:144) has written that comprehension of how education’s institutional framework informs individual behaviors and how organizations put these rules of conduct to use in certain organizational settings doesn’t necessarily mean that it (institutional forces) influences them in the same manner (italics added for emphasis)”. The institutional logics conflict over the ‘correct’ interpretation of SEI metrics is somewhat emblematic of the French nouvelle cuisine social movement of the 1970s that created “identity-discrepant cues” between classical French chefs and an emerging cadre of noveau chefs who wanted to replace the profession’s isomorphic practices (Rao, Monin, and Durand 2003:797). Similar to the noveau cuisine movement’s to establish new role identities amid a burgeoning class of chefs, today’s educational system, faculty, and students find themselves struggling to comprehend their new roles embedded in a neoliberal market framework (Bell 1976; Spillman 1999; Bartley 2005).

INSTITUTIONAL LOGICS FRAMEWORK

Although the two institutional logics informing the scholarly and market-based approaches to education may appear more similar than different on the surface, substantive differences are attributable to differences in teacher effectiveness ratings. To overcome this seeming antinomy, we propose that what differentiates the fields of interaction—besides their physical, environmental differences—are their competing institutional logics that embedded actors rely on to make sense and ‘manage’ their social identities, their social relationships (DiMaggio 1988; Nigam and Ocasio 2010; Scaraboto and Fischer 2013).
This is evident in recent reports that demonstrate that higher education administrators perceive online classes provided just as valuable a learning experience as traditional classroom settings in contrast to faculty who view online instruction as undermining academic progress (Allen and Seaman 2015). This near parody of views is not coincidental. For our analysis, we accentuate these differences in each model’s institutional logics by focusing on the online, transactional, and time sensitive structure of a market-based model in contrast to the scholarly model that embodies the traditional cultural autonomy of the academy and its time-dependent path to accomplishment.

WEBER’S IDEAL TYPES: THE PHOENIX AND PRINCETON MODELS OF HIGHER EDUCATION

To investigate the effect of the recent shift among competing institutional logics on student-instructor-administrator relations, we borrowed from Max Weber’s comparative model of ideal types to examine for substantive influences on the outcomes of SEIs in different institutional frameworks (Gerth and Mils 1948). To distinguish between the two competing logics—scholarly and neoliberal—we constructed two abstract models represented by the University of Phoenix (UOP) and Princeton University. Borrowing from Cottom (2017), we designated the UOP institutional logic model to the Lower Ed education field and the Princeton institutional logic to the elite higher education institutions generally out of reach of working class adults.

The ‘Phoenix’ model establishes students as instrumental actors searching to achieve higher education credentials by choosing market efficiencies and technological advantages over substantive traditional learning pedagogy. The ‘Princeton’ model in contrast, structures students’ behavior along a scholarly trajectory that prizes dominant cultural themes associated with distinction and credentialing that result in significant mentor relations and higher-income occupational trajectories (Bourdieu 1984). Thus, rather than view the interpretability of SEIs as simply informed by an isomorphic logic applied similarly across educational organizations, we begin with the notion that administrators, faculty, students, and parents are embedded within two institutional fields that are in competition for organizational dominance (Thornton, Ocasio, and Lounsbury 2012).

The Princeton model signifies the embedded institutional logic of faculty autonomy, shared governance, tenure, freedom of expression in the classroom, course quality over quantity, reliance on alumni and foundations for economic support, and large capital outlays for buildings and maintenance. The Princeton model exemplifies the relatively steady supply of high school graduates based on population demographics and the western European motivation for upward mobility.
The Phoenix model in contrast, relies on low-cost capital loans or public stock offerings, minimal buildings and maintenance outlays for virtual campuses and storefront campuses and classrooms, and categorizes instructors as human capital assets for accounting purposes. Additionally, the neoliberal logic in place constrains academic freedom by recasting professors as contract workers with minimal benefits and little bargaining power and autonomy (Tierney and Lechuga 2010; Mullin 2014; Lucal 2015; Olssen 2016).

A Phoenix-type college’s degree is a marketable asset, nothing more, nothing less. The Phoenix logic embodies neoliberal practices that accelerate specialization, encroaching on diversity of courses and schedules through limiting time commitment and the use of virtual reality that allows a compression of ‘where’ and ‘when’ class participation occurs with neither necessarily tied to the other (Cronin and Bachorz 2005; McPherson and Bacow 2015). The Phoenix model has lower student tuition and fees, shortened class time devoted to coursework, and accelerated class schedules, including concierge services that bolster revenues through student retention (Nickolai, Hoffman, and Trautner 2012). The market-based model maximizes student recruitment by marketing to an underserved precariat workforce (Standing 2011).

In summary, the multiple or coevolving institutional logics in the Princeton and Phoenix models highlights a widespread conundrum for administrators, faculty, students, and parents: which logic is informing actors’ social identity under changing inter-institutional relationships.

METHODOLOGY

Our study provides an instrumental case study of the inter-institutional logics that inform postsecondary public education SEI assessments in F2F and online settings. Expanding on prior research into SEIs, we are concerned with examining the discursive condition of students’ rating criteria as exemplified in the evaluation statements they rely on to make their judgments about instructor effectiveness. We take an abductive, inferential approach to deconstruct our sampled SEIs into signifiers of students and instructors’ institutional logics that inform their decision-making in their respective classroom (Tavory and Timmermans 2014).

As Thornton, Ocasio, and Lounsbury (2012) explained, a key theoretical advantage of institutional logics is its ability to expand institutional analysis by identifying subsystems, those inter-institutional mechanisms that re-introduce the relational context of embedded social interaction into the study of organizational behaviors. Theoretically, an institutional logic framework identifies those processes that inform actors’ decision-making heuristics under stable or uncertain conditions. As a sociological lens, an institutional logics perspective allows researchers to identify abrupt or slow changes in seemingly isomorphic conditions (organizational rules) and help
explain how actors have integrated, adapted, sustained, or changed their social positions under dominant field conditions (Di Maggio 1988; Susen 2007:37).

Institutional logics then, re-focuses research interest away from the antinomies of structure and agency, by identifying and helping to clarify the rules of play actors rely on to navigate social fields (Swidler 1986; Townley 1997; Wolfer and Johnson 2003; Chung and Luo 2008; Boxenbaum and Jonsson 2008). To test whether competing institutional logics affected SEIs ratings variations across classroom settings, we posited a latent factor variable – teaching effectiveness – that signified the ongoing social construction and sensemaking processes informing actors’ behaviors in F2F and online classes, ceteris paribus.

Our two models, Phoenix and Princeton, were partitioned by 1) SEI rating variables for each type classroom, and 2) instructor gender. The sample consists of instructors who taught in both a traditional and online class setting. The endogenous variables—SEI ratings—make up our latent factor, teacher effectiveness (Campbell et al. 2005; Lannutti and Strauman 2006; Steiner et al. 2006; Kline 2011). The selected key ratings criterion correlations, means, standard deviations, and correlations are provided in Table 1.

Using a structural equation factor model identified the direct and indirect effects among the SEI variables while flexibly taking into account measurement errors (Raykov and Marcoulides 2006) and allowed identifying the institutional logic processes informing students’ ratings of instructors in the two qualitatively different classes while holding the instructor variable constant. By controlling for instructors who taught both type of classes and received an evaluation in both class settings, we were able to make comparisons across their SEI ratings since the same SEI rating form was used by students in both class types. This further allowed accounting for our variables’ mediation linked to the presumed causal variables’ effects on our outcome latent factor of teaching effectiveness (Scheff 2006; Okrent 2007; Fiedler and Wänke 2009; Cloutier and Langley 2013; Thompson and Green 2013).

Model Specification

To examine for shifting institutional logics on students’ perceptions, a multilevel structural equation model with a latent factor was proposed as shown in Figure 1 (Powers 2012). The model allowed testing for invariance among estimated parameters between the two undergraduate class settings and instructors’ gender (Licht 1995; Thompson and Green 2013). Factor invariance tests provide for any number of covariates to be observed through the hypothesized latent factor represented by fitting “matrices of interrelationship indices – that is, covariance or correlation matrices – between all pairs of observed variables”, including testing for group means where appropriate (Kaplan 2009; Raykov and Marcoulides 2006:2).
STUDY DATA

In this study, we selected seven summary factors making up the nationally used IDEA evaluation form that measures student satisfaction, student learning, and students’ perceptions of instructor practices. The target variables in this study are instructors’ evaluation scores on six variables, a control criteria, gender (dummy coded 1 = female, 0 = male), and type of classroom setting (dummy coded 1 = traditional; 0 = online).

The six criteria were scored from 1 = non-effective teaching to 5 = most effective teaching for the following shortened statements: (1) teaches course fundamentals, (2) inspires/challenges students to do more than basic class work, (3) introduces stimulating ideas, (4) develops rapport with students, (5) provides sufficient and timely feedback to students, and (6) encourages student involvement in their course learning. The inclusion of the seventh indicator as a bias control variable – ‘I really wanted to take this course regardless who taught it’—was intended to determine if students’ ratings were influenced based on prior knowledge of the instructor of their course.

The data set comprised instructors’ summary IDEA evaluation scores who taught in both a face-to-face class (N = 539) and virtual classroom environment (N = 166) within their respective academic departments. The sampling frame included the undergraduate semester periods Fall 2010, Spring 2011, Fall 2011, and Spring 2012 (excluding summer sessions). Given prior research into brick-and-mortar and online venues, we theorized teacher effectiveness as the outcome latent factor of student’s relevant decision-making heuristics based on the six key evaluative criteria taken from the IDEA student evaluation form (means are provided in Table 1 above for the two models). We hypothesized that our latent factor would parsimoniously identify covariations among the evaluation criterion scores, fit our specified confirmatory models, and allow for identifying the institutional logic processes informing students’ perceptions of instructor effectiveness by type of class setting. The latent factor model also accounts for nonnormality in the sampled groups (Brown 2006).

The sampling unit approximates U.S. colleges that have used an established IDEA evaluation system of faculty effectiveness ratings. Although Frankfort-Nachmias (2008) has asserted that generalizations to a population must use a probability sample, costs and faculty autonomy concerns limited using a randomized sampling design. Instead, the sampling frame represented

---

1 The IDEA Student Ratings System originated at Kansas State University in the 1968. The evaluation criteria was designed to provide a metric focusing on improving instructors’ pedagogical styles. With help from the W. K. Kellogg Foundation, The IDEA Center was established in 1975 and the IDEA Student Ratings system was made available to other colleges and universities. The IDEA Center is a nonprofit entity and its evaluation metrics are widely used throughout the United States for a variety of instructor and administrator purposes in higher educational settings (IDEA 2015).
all course evaluations from both type class environments \((N_{\text{Traditional + Online}} = 705)\) partitioned by instructor gender \((N_{\text{male}} = 355; N_{\text{female}} = 350)\) obtained from a college of humanities and social sciences at a midsized public university in Texas. To control for potential data dependency due to the nesting structure (i.e., students were nested within instructors they rated), we analyzed the data with the Type = Complex Routine in Mplus program (Version 7.2; Muthen and Muthen 1998-2014), which directly accounted for dependence among the sample during data analyses.

Our statistical purpose was to compare each of our six (seven if counting the bias control variable) criterion variables used in the IDEA evaluation statements that students selected to rate their instructors. Additionally, we then partitioned the criterion variables with our teaching effectiveness latent factor, controlling for instructor gender and type class setting (traditional or online). Moreover, a control variable measuring student bias (I would have taken this course regardless of familiarity with the instructor) was included\(^2\).

Research Hypotheses

Guiding our research were the following hypotheses:

\(H_1:\) The seven endogenous variables’ coefficients explain a significant amount of the latent variable’s measurement variance and provide verification that the latent factor (teaching effectiveness) and other parameters are invariant, regardless the institutional field and subsequent institutional logics.

\(H_2:\) Students’ assessments of teaching effectiveness between the two institutional fields and associated logics were invariant by instructor gender.

\(H_3:\) Students’ assessments of teaching effectiveness were not differentially affected by competing institutional logics between institutional fields across classroom settings.

ANALYSES & RESULTS

The factor model illustrated in Figure 1 is a conceptualization of students’ ratings effect on the latent variable, teaching effectiveness by

\(^2\) Although some instructors received more than one evaluation per type of classroom setting, we control for inter- and intra-variability in the obtained endogenous and exogenous variables by restricting generalizations to combined variable effects on the evaluation criterion.
institutional setting—Phoenix or Princeton models. Given the multilevel data structure (with 706 students nested within 61 courses with an average of 11.57 students per course) and possible data dependency, we first obtained the intra-class correlation (ICC) for each variable to determine the multilevel model’s appropriateness (Lai and Kwok, 2015). As presented in Table 1, the ICCs of the seven variables ranged from .15 to .43, which indicated substantial non-independency in our data, and the need to account adequately for dependency outcomes in the analyses. The means, standard deviations, and the zero-order correlations of the six variables are also presented in Table 1 (the control variable was not significant in any of our models and is not shown in the table).

Because we were interested in examining group differences between our six criterion variables, instructor gender, and type of classroom, we relied on hierarchical measurements of invariance—configural, metric, and scalar—to confirm the fit of our hypothesized factor model and test for significant differences of loadings between our institutional logics class environments, including latent factor intercepts on the six SEI variables investigated.

Conventional use of a hierarchical invariance tests ensures that parameter estimates and assumptions ‘best’ reflect model parameters at each step in the estimating process before attempting to compare the latent variable intercepts, the measurement equation intercepts, and the means of the latent exogenous variables. Bollen (1989:366) writes, “At a minimum the invariance of form and factor loadings should hold before testing restrictions on means and intercepts”. The invariance test results confirmed our model was appropriate and as per convention, allowed for further inductive investigation into group differences (type of classroom and gender) as shown in Tables 2 and 3.

MODEL TESTING

Hypothesis 1: Testing the Multilevel Latent Factor Model Fit

As a necessary first step, we tested the null hypothesis that our sample population covariance matrices were equal. Failure to reject the null hypothesis would imply the equality of population covariance matrices was plausible, which in turn indicates that equality of the factor loadings, the matrix’s unique variances, and factor variances and covariances were not significantly different based on the teaching effectiveness latent factor with the seven selected (endogenous) indicators.

Our statistical model fit the data adequately $\chi^2 (21) = 93.03, \ p < 0.001; \ \text{RMSEA}=.07, \ \text{SRMR}=.06, \ \text{and CFI}=.95$, allowing for tentative acceptance that the seven SEI variables’ coefficients explained a significant amount of the latent variable’s measurement variance. This would provide verification that the latent factor (reaching effectiveness) and other parameters were invariant,
regardless the institutional field and subsequent institutional logics. However, when we examined the parameter estimates, the loading of the student bias indicator was not significant ($\lambda = .08, p = .08$). To further test our model fit, we examined the same seven-indicator model by different gender groups and classroom delivery methods and found the same non-significant factor loading across groups for the bias variable. Hence, we excluded the control variable from further analyses and reexamined the model with our six indicator variables. The model produced an adequate fit ($\chi^2 (15) = 76.79, p < 0.001$; RMSEA = 0.08, SRMR = 0.06, & CFI = 0.95) with all parameter estimates statistically significant ($p < 0.05$).

The model fit analyses of the six indicators significantly loaded on the teaching effectiveness latent factor (with standardized factor loadings ranging from .72 to .90 and the corresponding explained variances [or R$^2$] ranging from .52 to .81). The measurement invariance findings confirmed that the multilevel factor model provided an adequate estimation and fit of institutional processes among factor loadings, including our teaching effectiveness latent variable, i.e., teaching effectiveness.

**Hypothesis 2: Testing Gender Differences**

We conducted the factorial invariance test following recommended procedural steps (Millsap 2011): The configural invariance test examined whether the same factor model (with all six indicators loaded on the teaching effectiveness factor) adequately fit by gender, regardless the institutional setting. We found this model fit marginally well to the data; based on the modification index, we correlated the residuals between the inspired students and the student-involvement variables. This modified model produced an adequate fit ($\chi^2 (16) = 36.80, p < 0.05$; RMSEA = 0.06, SRMR = 0.02, and CFI = 0.99).

We then calculated a metric invariance test to examine whether the relations (i.e., the factor loadings) between the six indicators and the teaching effectiveness latent factor were the same across institutional settings regardless of gender. The non-significant chi-square difference test ($\Delta \chi^2 (5) = 9.81, p > 0.05$) indicated that the six factor loadings were equal for both institutional settings. In other words, the students’ ratings on the six SEI indicators and the teaching effectiveness latent factor were not different based on an instructor’s gender.

As a final invariance measurement test, we calculated a scalar metric by constraining both factor loadings and latent intercepts of the criterion indicators to be equal for both men and women as given in Table 2. The non-significant chi-square difference test (i.e., $\Delta \chi^2 (5) = 8.30, p > 0.05$) confirmed that the factor loadings and latent intercepts were equal for instructors in both classroom settings, irrespective of gender. Based on the scalar findings, we
then examined for gender difference on the mean of the teaching effectiveness latent factor; the results confirmed our earlier findings that there were no significant differences exhibited by students’ ratings in either class based on an instructor’s gender ($\Delta \mu_{\text{teaching_effectiveness}} = \mu_{\text{teaching_effectiveness_Female}} - \mu_{\text{teaching_effectiveness_Male}} = 0.11, p \leq 0.12$).

**Hypothesis 3: Teaching Effectiveness and Competing Logics in F2F and Online Setting**

We repeated the invariance test steps to examine for possible institutional differences between classrooms on the teaching effectiveness latent factor. By including two additional correlated residuals, namely, *inspired* and *student-involvement*, and between *teaches fundamentals* and *provides stimulating learning environment*, we significantly increased our model fit ($\chi^2(14) = 32.83, p < 0.05; \text{RMSEA} = 0.06, \text{SRMR} = 0.02$, and $\text{CFI} = 0.99$).

The metric invariance test confirmed our suspicions that two of the factor loadings “were ‘statistically not invariant’” by institutional setting. To obtain a better fitting model, we constrained four factor loadings to be equal across the classroom types while freely estimating the factor loadings of *inspired students to do more* and *teaches fundamentals*. As presented in Table 4, this partial metric fit the data well ($\chi^2(17) = 36.34, p < 0.05; \text{RMSEA} = 0.06, \text{SRMR} = 0.06$, & $\text{CFI} = 0.99$).

To expand our hypothesis test that students’ ratings of instructor effectiveness were not affected by competing institutional logics across classroom settings, a chi-square difference test was used as given in Table 3. The results demonstrated that with the exception of the factor loadings for *inspired students to do more* and *teach fundamentals*, the remaining four-factor loadings were invariant between the institutional settings. In other words, the factor loadings for *inspired* and *fundamentals* demonstrated a stronger relation to the teaching effectiveness outcome in our hypothesized Princeton model compared to the Phoenix model.

We then calculated the partial scalar metric based on our invariance model and again found that two (out of the six) latent intercepts were significantly different. By constraining four of the latent intercepts to be equal between the two theorized institutional fields while freely estimating the latent intercepts of the *stimulates learning* and develops *close rapport* indicators, the partial scalar invariance model’s fit improved substantially ($\chi^2(20) = 39.21, p < 0.05; \text{RMSEA} = 0.05, \text{SRMR} = 0.06$, and $\text{CFI} = 0.99$).

The non-significant chi-square test (i.e., $\Delta \chi^2(3) = 1.87, p > 0.05$) reconfirmed our earlier decision to constrain the latent intercepts in order to more closely examine for invariance among SEI variables between the two
institutional settings. Given the overall findings of the partial scalar metric, we further tested the potential mean difference of our teaching effectiveness latent factor. The results confirmed our institutional logics model that instructors in the Princeton ideal type received a significantly higher mean teaching effectiveness rating than when they were embedded in the Phoenix model ($\Delta \mu_{\text{teaching\_effectiveness}} = \mu_{\text{teaching\_effectiveness\_Traditional}} - \mu_{\text{teaching\_effectiveness\_Online}} = .19, \ p<.01$).

DISCUSSION

The multilevel latent-factor model provided significant statistical flexibility to examine simultaneously direct and indirect institutional effects on students’ ratings of teaching effectiveness via the selected criterion variables. The latent factor model also accounted for expected measurement errors in the data across the two analytical models (Raykov and Marcoulides 2006). The goodness of fit results further confirmed that the estimated parameters best-replicated students’ ratings used to measure teacher effectiveness in the two idealized institutional models. In addition, the latent factor invariance tests demonstrated that students’ apperceptions were affected by different logics—Princeton versus Phoenix—based on their embeddedness in a particular field of social interaction, introducing a new construct for future investigation (Driscoll et al. 2012). Importantly, our case study revealed that students’ assessments of instructors were not biased by instructor gender; a significant finding given that a substantial number of prior studies have reported students’ gender biases effecting instructor assessments (Beqiri, Chase, and Bishka 2010; Benton and Cashin 2012; Benton and Li 2015).

Overall, the statistical findings and model results confirmed that the SEI variables – inspired and fundamentals were rated higher in the Princeton model compared to the Phoenix, online classroom. We inferred from this that students interpreted pedagogy styles significantly different based on their inter-institutional contexts. Because inspired students was defined as motivating students to increase their participation in class materials in and outside class, the lower ratings in an online setting also supports previous research that having students increase their class work in a Phoenix-type institutional setting was considered unjust by students because of the additional time commitments (Klaus and Chagchit 2009). Conversely, inspiring students in F2F classes required an institutional logic substantively different from an instrumental communicative platform found in a virtual, asynchronous setting. For example, opportunities to confer with other students, to negotiate and clarify with instructors over principles of classwork rigor, fairness, and instructional quality are important socio-performative characteristics in F2F classes that make up instructor-student fields of interaction (Gregory 2012; Pardasani et al. 2012).
The factor loadings for *teaches fundamentals* was also significantly higher in the Princeton model, further demonstrating a strong association with the physical, performative aspect found in F2F classrooms because of students’ overall educational socialization from grade school to college (Howland Moore 2002; Rovai et al. 2006; Carillo 2007; Nichols 2011). However, in the Phoenix model, the quality of student social interaction in terms of self-assessment does not indicate the same field relations found in F2F classes (Howland and Moore 2002; Robinson and Hullinger 2008).

Because prior research has shown that an absence of social cues, latency of response times and interactions, and time allotted for social exchanges in an online platform undermined interpersonal relations (Okdie and Guadagno 2008), we posit that the analyses provided empirical support of our hypothesized ideal model types. For example, in their work, Xie, DeBacker, and Ferguson (2006) found that students in a virtual classroom were unlikely to develop and demonstrate engaged learning behaviors that indicated student-to-student learning as exemplified in F2F classes when instructors had failed to develop and ‘post’ a class compliance rubric in online settings (Novak 2011).

Moreover, the point estimates (intercepts) observed for the variables—*stimulating* and *rapport*—in the Princeton model demonstrated that the isomorphic force of the scholarly institutional setting generally began with a higher rating. This finding is reasonable given the longer socialization experience of students embedded in the scholarly, public institutions that support traditional pedagogy. This further accentuates the comparative differences between the two competing institutional logics that are currently causing conflict in the academic and public spheres (Tu and McIssac 2002). For example, instructors and administrators are at odds over student ratings in different types of classrooms—synchronous and asynchronous—because they have not recognized that there are at least two competing institutional logics informing students’ perceptions of teaching effectiveness given the classroom or course’s organizational setting.

**CONCLUSION**

First, our case study demonstrated that course criteria and instructor pedagogy alone do not explain student classroom preferences and assessments of teaching effectiveness. By partitioning actors’ informed behaviors from the sampled SEIs into two models—Princeton and Phoenix—we were able to identify differences in instructors’ assessments by competing institutional logics. Overall, the key variables—inspired, fundamentals, stimulating and rapport—indicated that it is not the evaluation criteria that are ‘good’ or ‘bad’ measures of teaching effectiveness, per se; but it does seem reasonable to conclude that the inter-institutional logic affects differently stakeholders’ sense making heuristics where the act of interpreting the ‘agreeableness of
practices’ emerges. Because SEIs outcomes were significantly different given the field context, our study offered a theoretical framework that included the inter-institutional mechanisms and processes affecting administrator-faculty-student interactions and ultimately, addressed the growing crisis in the academy related to the ascendance of neoliberal market-based principles in university settings.

Inductively, our models demonstrated institutional logics do matter in an evaluative context. One key finding that emerged demonstrated that SEI assessments were not biased by faculty gender. This is a welcome finding because it indicates that students’ gender-biases may have shifted, as gender stereotypes have been debunked. We hope other researchers will be able to replicate this finding in the future. Broadly, we asked, was a shift in educational authority affecting students and teachers apperceptions, their value-laden classifications (substantive and instrumental) that acted as a filter and allowed them to comprehend what approximates effective learning. We believe our case study models demonstrated that it does.

Thus, it would seem reasonable to recommend that teaching effectiveness metrics might better serve administrators, faculty, students, and parents if SEI ratings assessments were constructed based on an understanding that the organizational setting students and instructors are embedded were informed by qualitatively different institutional logics. We suspect that evaluation schemes that purport a ‘one-size fits all’ mentality may be at the heart of today’s SEI conflicts precisely because of this oversight in recognizing the institutional logics influence over actors’ apperceptions and subsequently, their behaviors.

FUTURE RESEARCH AND LIMITATIONS

The case study model combined with a multilevel latent factor model demonstrated a robust analytical approach to the complexity of identifying inter-institutional processes and outcomes regarding teaching effectiveness metrics. We believe that additional case studies that examine for variations in institutional logics within and across educational settings and sectors offers a nascent approach and would bolster efforts to arrive at an acceptable metric of teaching effectiveness that both benefited faculty, administrators, and students over the long term. Because the sample collected was not racially or ethnically diverse, future research should attempt to address this methodological weakness. Additionally, the SEIs sampled were for either traditional or online courses offered and did not consider SEIs for blended or ‘flipped’ class platforms. Another area unexamined was the Massive Open Online Courses and students’ evaluations under those institutional conditions. Finally, because of student privacy guidelines, we were unable to identify which students in the pool of SEIs collected had taken courses in both types of classroom settings. Thus, our results cannot be generalized to students particularly or to
undergraduate colleges in the United States given the sampling limitations of our collection method. We believe that addressing this weakness in our matching of instructors-students-and classroom settings would provide a more transparent test of our institutional logics theoretical framework and offer policymakers a more transparent tool for constructing teacher effectiveness instruments.
REFERENCES


Dunn, Mary B. and Candace Jones. 2010. Institutional Logics and Institutional Pluralism: The Contestation of Care and Science Logics in


Howland, Jane L., and Joi L. Moore. 2002. “Student Perceptions as Distance Learners in Internet-Based Courses.” Distance Education 23(2): 183-195.


Young, Suzanne and Heather E. Duncan. 2014. “‘Online and Face-to-Face Teaching: How Do Student Ratings Differ?’” *MERLOT Journal of Online Learning and Teaching* 10(1): 70-79.
<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Rapport</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Student involved</td>
<td>.05</td>
<td>.63**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Stimulating</td>
<td>.12*</td>
<td>.81**</td>
<td>.65**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Inspired</td>
<td>.10*</td>
<td>.78**</td>
<td>.75**</td>
<td>.77**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Feedback</td>
<td>.03</td>
<td>.72**</td>
<td>.47**</td>
<td>.64**</td>
<td>.64**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Fundamentals</td>
<td>.06</td>
<td>.63**</td>
<td>.45**</td>
<td>.71**</td>
<td>.61**</td>
<td>.55**</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$M^\dagger$</th>
<th>$SD^{\ddagger}$</th>
<th>ICC$^{\ddagger\ddagger}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.14</td>
<td>.49</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>4.43</td>
<td>.47</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>4.08</td>
<td>.51</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>4.26</td>
<td>.47</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>3.96</td>
<td>.48</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>4.29</td>
<td>.52</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>4.14</td>
<td>.47</td>
<td>.29</td>
</tr>
</tbody>
</table>

*Note. * $p < .05$, ** $p < .001$

$M^\dagger$: Mean; $SD^{\ddagger}$: Standard Deviation; ICC$^{\ddagger\ddagger}$: Intra-Class Correlation
Table 2. Results for Measurement Invariance Tests by Gender

<table>
<thead>
<tr>
<th>Measurement Invariance Model</th>
<th>Configural</th>
<th>Metric</th>
<th>Scalar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Loading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapport</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Student Involved</td>
<td>.775</td>
<td>.903</td>
<td>.835</td>
</tr>
<tr>
<td>Stimulating</td>
<td>.966</td>
<td>1.095</td>
<td>1.035</td>
</tr>
<tr>
<td>Inspired</td>
<td>.891</td>
<td>1.079</td>
<td>.983</td>
</tr>
<tr>
<td>Feedback</td>
<td>.810</td>
<td>1.042</td>
<td>.931</td>
</tr>
<tr>
<td>Fundamentals</td>
<td>.725</td>
<td>.959</td>
<td>.861</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Involved</td>
<td>3.998</td>
<td>4.156</td>
<td>3.998</td>
</tr>
<tr>
<td>Stimulating</td>
<td>4.223</td>
<td>4.292</td>
<td>4.223</td>
</tr>
<tr>
<td>Inspired</td>
<td>3.890</td>
<td>4.029</td>
<td>3.890</td>
</tr>
<tr>
<td>Feedback</td>
<td>4.208</td>
<td>4.375</td>
<td>4.208</td>
</tr>
<tr>
<td>Fundamentals</td>
<td>4.128</td>
<td>4.145</td>
<td>4.128</td>
</tr>
<tr>
<td>Fit Statistic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$ (df)</td>
<td>36.801* (16)</td>
<td>46.648* (21)</td>
<td>53.243* (26)</td>
</tr>
<tr>
<td>$\Delta \chi^2$ (df)</td>
<td>9.814 (5)</td>
<td>.059</td>
<td>.054</td>
</tr>
<tr>
<td>RMSEA</td>
<td>.061</td>
<td>.129</td>
<td>.137</td>
</tr>
<tr>
<td>SRMR</td>
<td>.994</td>
<td>.992</td>
<td>.992</td>
</tr>
<tr>
<td>CFI</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05
Table 3. Results for Measurement Invariance Tests by Classes

<table>
<thead>
<tr>
<th></th>
<th>Measurement Invariance Model</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Configural</td>
<td>Metric</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Online</td>
<td>Traditio</td>
<td>Online</td>
<td>Traditio</td>
</tr>
<tr>
<td>Loading</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulating</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Rapport</td>
<td>.988</td>
<td>1.011</td>
<td>1.001</td>
<td>.999</td>
</tr>
<tr>
<td>Student Involved</td>
<td>.913</td>
<td>.869</td>
<td>.894</td>
<td>.888</td>
</tr>
<tr>
<td>Inspired</td>
<td>.923</td>
<td>1.119</td>
<td>.930</td>
<td>1.119</td>
</tr>
<tr>
<td>Feedback</td>
<td>.890</td>
<td>1.028</td>
<td>.975</td>
<td>.975</td>
</tr>
<tr>
<td>Fundamentals</td>
<td>.582</td>
<td>.956</td>
<td>.586</td>
<td>.951</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulating</td>
<td>4.022</td>
<td>4.330</td>
<td>4.022</td>
<td>4.330</td>
</tr>
<tr>
<td>Student Involved</td>
<td>3.968</td>
<td>4.110</td>
<td>3.968</td>
<td>4.110</td>
</tr>
<tr>
<td>Inspired</td>
<td>3.825</td>
<td>4.00</td>
<td>3.825</td>
<td>4.00</td>
</tr>
<tr>
<td>Feedback</td>
<td>4.158</td>
<td>4.332</td>
<td>4.158</td>
<td>4.332</td>
</tr>
<tr>
<td>Fundamentals</td>
<td>3.980</td>
<td>4.185</td>
<td>3.980</td>
<td>4.185</td>
</tr>
<tr>
<td>Fit Statistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>32.829* (14)</td>
<td>36.344* (17)</td>
<td>39.206* (20)</td>
<td></td>
</tr>
<tr>
<td>$\Delta \chi^2$</td>
<td>2.494 (3)</td>
<td>1.872 (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEA</td>
<td>.062</td>
<td>.057</td>
<td>.052</td>
<td></td>
</tr>
<tr>
<td>SRMR</td>
<td>.020</td>
<td>.055</td>
<td>.057</td>
<td></td>
</tr>
<tr>
<td>CFI</td>
<td>.990</td>
<td>.989</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05
Teaching Effectiveness

![Diagram showing the relationship between Bias, Rapport, Student Involved, Stimulating, Inspired, Feedback, and Fundamentals leading to Effect]

Figure 1