ALIGNING TEACHER PREPARATION, PROFESSIONAL DEVELOPMENT AND EVALUATION: THE ORTHODOXY OF TAP TEACHERS AND TEACHING

JESSICA HOLLOWAY

Abstract

This paper looks at one of the most popular teacher accountability systems in the USA, TAP: The System for Teacher and Student Advancement, which is a comprehensive program that meets the requirements of multiple large federal grants schemes. The paper shows that, after consortia in Arizona and Texas won multimillion-dollar grant awards, thousands of teachers in these states have become engrossed within a TAP circuit that shapes multiple domains of their professional identity and growth. TAP's ubiquitous presence helps build system alignment, but it also means that teachers are increasingly limited in their exposure to alternative philosophies, practices, or measurements of teaching. Therefore, I challenge the assumptions of alignment and illustrate how such alignment flattens and overly-simplifies the plurality and complexity of teaching, even though present times require adaptability within schools. What I ultimately argue is that the precise and comprehensive alignment of this system enforces and reinforces TAP as the orthodox reality of what it means to be and become a teacher.

Keywords

teacher accountability, policy alignment, TAP System, rubrics, value-added measurement

Introduction

The past three decades have brought about a significant increase in standards-based and data-driven accountability systems within education (Hardy, 2019; Lingard & Rizvi, 2010; Savage & Lewis, 2018). Teachers, in particular, have been subjected to high-stakes accountability that rely on calculating tools (e.g., value-added models, performance rubrics) and punitive actions (e.g., merit pay, termination) that have fundamentally reshaped teacher subjectivities in the image of data (Ball, 2003, 2015; Holloway & Brass, 2018). For example, teachers face increased pressures to rely on standards, test data and evaluative instruments (e.g., rubrics) to guide their classroom practices (Bradbury 2019; Bradbury & Roberts-Holmes; 2017; Garver, 2019; Hardy, 2019), which is reinforced in three interrelated domains of practice: teacher preparation, professional development (PD), and performance appraisal (see Holloway, 2019). As the motivation for securing standards and data as the core of these domains grows increasingly prevalent, the alignment between them becomes equally important. In other words, if teachers are to be evaluated on the standards and their ability to produce good data, then it can be expected that their training and development experiences prepare (or discipline) them for this type of examination (cf. Foucault, 1977; see also Ball, 2003; Perryman, 2009).

Accordingly, in the United States, alignment between teacher preparation, professional development, and performance appraisal has been a key priority of large federal grant schemes that incentivize accountability reform. Two specific grant programs—the Teacher Incentive Fund (TIF) and the Supporting Effective Educator Development (SEED) schemes—have been particularly influential in procuring alignment between various domains of teaching practice. While pursuits to increase alignment between training, development and evaluation might help prepare teachers for the data- and standards-based environments they will inevitably face, there are critical questions regarding the potential consequences of alignment that must also be considered. Drawing on recent work by Savage and O'Connor (2019) that theorises around the "problem with policy alignment," I use this paper to critically examine how strict alignment priorities within US teacher policy have created a sort of orthodoxy of teachers and teaching practice. The empirical material for the paper derives from publicly available documents and artefacts associated with two grant-funded projects that have supported the purchase and implementation of a performance-based accountability system across two subnational systems within Texas and Arizona. The grant projects—the *Arizona Ready for Rigor* project and *TAP Connect* (in Texas)—provide interesting cases for considering the consequences of alignment, as both of these projects have established partnerships between consortia of

K-12 schools, university-based teacher preparation programs and the non-profit organisation, the *National Institute for Excellence in Teaching* (NIET). Through these partnerships, a single accountability system, *TAP: The System for Teacher and Student Advancement* (formerly known as the Teacher Advancement Program and hereafter referred to as the TAP System), has been embedded within teacher preparation programs, K-12 professional development strategies and evaluation systems across both states.

I aim to show that, after the two consortia won multimillion-dollar grants, a vast number of teachers in these states have been engrossed within a TAP circuit that shapes multiple domains of their professional identity and growth. The grants secured the TAP System as the teacher preparation framework, teacher evaluation system and professional development program in 58 Arizona schools and 27 Texas schools, as well as one Arizona university and one Texas university. As such, thousands of teachers within these states have transitioned directly from TAP-based preparation programs into TAP-based schools that use TAP evaluation, professional development, and personnel incentive systems (e.g., merit-based pay, promotion).

Ultimately, I challenge the assumptions of alignment and illustrate how such alignment flattens and overly-simplifies the plurality and complexity of teaching, even though present times require adaptability within schools (e.g., schools needing to respond to demographic changes).

I use policy documents, promotional materials, and accountability instruments and artefacts (e.g., grant applications, press releases, TAP and teacher preparation websites, professional development handbooks, observation rubrics) to (1) map TAP's presence in the interrelated domains of teacher preparation, professional development, and evaluation, (2) identify key features of and relationships between these domains of practice, and (3) use tenets of pluralism (Connolly, 2005) to problematise goals that prioritise strict alignment of policy, practice and institutions. What I argue is that the precise and comprehensive alignment of this system enforces and reinforces TAP as the orthodox reality of what it means to be and become a teacher.

The paper unfolds in the following ways: first, I provide an overview of the policy context, while trying to avoid unnecessarily complicating the argument with excessive jargon and acronyms (see footnotes for additional information where interested). Then, I move to describe my analytical approach to the paper. Next, I provide the empirical analysis to illustrate how the TAP orthodoxy creates a fixed ontological space for the teacher becoming and being. This is where I develop the concept of *orthodox ontology*, while drawing on Foucault's (1980) view of discourse and Connolly's (2005; 2013) political theory work on pluralism. I conclude with a discussion about why plurality and complexity should be embraced when it comes to teachers and teaching in contemporary times.

Policy Context

Over the past few decades, a series of federal grant programs have incentivized most US states and school districts to radically modify their teacher evaluation policies and practices to include new modes of numerically measuring and evaluating teacher quality (Collins & Amrein-Beardsley, 2014; Hewitt & Amrein-Beardsley, 2016). Specifically, these new accountability frameworks have had several effects, including: (1) increased reliance on numbers-based rubrics for measuring teacher performance; (2) introduced value-added measures (VAMs), which are statistical tools that directly link student test scores to teacher effects; and (3) compelled schools to increase the stakes attached to teacher evaluation outcomes. These grant programs sit within a broader movement of standards-based reform, which has characterized much of education policy in the US (and abroad) for more than 30 years. What has become increasingly emphasized in the standards-based reform movement is the motivation to align policies and practices between various schooling domains (Lee, 2019; Martone & Sireci, 2009; Savage & O'Connor, 2019). Common examples of alignment might include aligning standards and curriculum with student testing, colleges of teacher education with accreditation requirements, teacher evaluation with professional development programs, or alignment between state (or other subnational) education systems in the name of raising standards and ensuring equal opportunity (see, for example Savage & O'Connor, 2019). Through No Child Left Behind, Race to the Top and a number of other policy incentive programs, policy alignment has been prioritized as a means for improving educational quality and ensuring high standards for both teaching and learning.

For this paper, I turn my attention specifically to recent US grant schemes that have prioritized alignment between teacher preparation, teacher evaluation, professional development, and personnel systems. The empirical focus of the paper is on two large partnerships that have won Teacher Incentive Fund (TIF) and Supporting Effective Educator Development (SEED) grant awards, both of which involved the National Institute for Excellence in Teaching (NIET). NIET presents a compelling case for examining alignment given its long history in partnering with schooling systems (both at the K-12 and university levels) to help schools build systematic infrastructure that purposefully aligns with federal- and state-level policy initiatives. Indeed, NIET's signature program, the TAP System, has been advertised specifically on the basis of its alignment with the federal TIF grant scheme, boasting that eight of the 34 awarded grants in 2010 were applicants who adopted the TAP System, NIET is a 501(c)(3) public charity that primarily works with high-need schools and has formed partnerships with a number of states, school districts, and universities. Their TAP System,

which is a comprehensive performance-based compensation system, was first used during the 2000–2001 academic school year (Daley & Kim, 2010; 2012), and has since been used in approximately 600 schools across 19 states (estimated to have affected more than a million students). The TAP System is composed of three measurements of teacher performance, which are combined for evaluative and comparative purposes: VAMs, classroom observation rubrics, and skills, knowledge and responsibility rubrics. VAMs are calculated once a year (by external statisticians), and classroom observation rubrics are used four times a year during formal observations. The rubric consists of 19 indicators, making up three constructs ("1. Instruction," "2. Designing and Planning Instruction," and "3. The Learning Environment"), and each domain is broken down into specific criteria for teachers to complete during each lesson. For example, "Instruction" is broken down into 12 indicators (e.g., Standards and Objectives, Motivating Students), which are also broken down into specific criteria for each of the categories of judgment (i.e., Exemplary, Proficient, Unsatisfactory). Before the observations, one-on-one pre-conferences are held to discuss the planned lesson. After the observations, post-conferences are held to discuss the observed lesson, rubric scores, and areas of refinement (i.e., to be improved) and reinforcement (i.e., to continue as is). Rubrics are also used to measure teachers' "skills, knowledge, and responsibilities," which concerns matters such as participation in professional development workshops and the teacher's ability to effectively self-evaluate. The final evaluation is a composition of rubric scores and VAM scores, which are then used to rank teachers and distribute performance-based bonus pay accordingly.

According to NIET, "The core of TAP lies in its instructional rubric of best practices, which university, school and district staff use as a foundation for their training and performance evaluations" (NIET, n.d., n.p.). As will become increasingly clear throughout this paper, there is little, if any, part of the teacher's training or professional experience that is not affected and shaped by the TAP System rubric. This has profound implications in terms of the material, philosophical and theoretical ideals to which pre-service and practicing teachers are exposed, which I will revisit in the analysis and discussion sections of the paper. Therefore, it is worth providing a bit more detail about the underlying ideology and main features the rubric. According to NIET, the TAP rubric was designed using the following resources (NIET, 2011):

- Interstate New Teacher Assessment and Support Consortium (INTASC)
- National Board for Professional Teacher Standards
- Massachusetts' Principles for Effective Teaching
- California's Standards for the Teaching Profession
- Connecticut's Beginning Educator Support Program
- New Teacher Center's Developmental Continuum of Teacher Abilities
- Danielson's Framework for Teaching

Charlotte Danielson—who joined NIET in 2007 as Distinguished Scholar—developed the *Framework for Teaching* (Danielson, 2007), which has become the most widely used teaching framework across US schools to date (The Danielson Group, n.d.). Her framework was also significantly influential in the development of the TAP rubric. The two frameworks are purported to be "research-based" (NIET, n.d., n.p.), though these claims have been questioned over the years (see Sloat, Amrein-Beardsley, & Sabo, 2017). They are both said to be based on a constructivist understanding of teaching and learning, but can also be conceptualised as a set of discreet and measurable activities and dispositions. As such, the rubric attempts to capture as many aspects of "effective teaching" as possible, so as to provide a common language for teachers and evaluators linking teacher practice with improved student achievement (NIET, n.d.). In doing so—I will argue throughout the rest of the paper—the rubric becomes cumbersomely prescriptive, especially given the significant priority it is given during the teachers' preparation, development, and evaluation. What I aim to show is that, regardless of how good or bad the rubric might be, it is the privileging of the TAP framework (at the expense of alternative approaches to teaching) that is problematic. I should stress here that the purpose of my argument is to problematise the alignment between teacher preparation, professional development, and evaluation, rather than comment explicitly on the quality of TAP or the ideological nature of the system. It is the idea that teachers' professional being and becoming are oriented to a single framework with which I seek to contend. First, though, I will describe my analytical approach to the paper.

Analytic Process

Broadly speaking, the paper is situated within the field of critical policy sociology, and therefore does not assume that the methodological approach, nor the associated conclusions, can be validated or free from bias. Rather, I work from the epistemological assumption that empirical material can help illustrate how discourses are shaping "conditions of possibility" for what can be said, thought, and done. Analytically, I drew on the work of Bacchi (2000) and Foucault (2002) to map TAP's presence and operation within and between the various domains of teaching practice. Assuming a policy-as-discourse framework (cf. Bacchi, 2000), I guided the analysis by questioning (1) how TAP defines the key features of and relationships between the domains, and (2) how strict alignment structures what is possible within the Arizona and Texas systems. To do this, I collected various policy materials related to the TAP grants (i.e., *Arizona Ready-for-Rigor*, *Planting the SEED*, and *TAP Connect*), such as grant applications and grant scheme requirements.

I also collected policy documents and artefacts related to the systems involved in the grant partnerships, such as university websites and teacher preparation program and course descriptions, TAP flyers, websites, videos, and press releases. I also collected data on the TAP System practices and instruments, including rubrics, student growth model information, evaluator forms, handbooks, websites, and videos.

I started with open coding (Saldaña, 2015) to reduce the data to a manageable and relevant collection of documents, artefacts, and other media materials. Then, I attended to the policies, practices, and instruments that served as the key features of each domain, before comparing the domains to see how they related to one another. I first analysed each state-based case (Arizona and Texas) separately to make comparisons between the cases without making generalized assumptions about both as a whole. The first part of this stage was descriptive in that I was interested in identifying the practices and instruments that were used in each of the domains. The second part of this stage was more analytical in that I was interested in understanding how the practices and instruments shaped the ontological space of teacher development and being. For this part, I leaned on Foucault (2000), who was interested in the relationship between discourses and the formation of subjects as determined (and re-determined over time) by registers of knowledge available. He wrote that "what we should do is show the historical construction of a subject through a discourse understood as consisting of a set of strategies which are part of social practices" (as cited in Davies & Bansel, 2010, p. 5-6). Important to note here is that the subject is not a passive individual who has discourses done to them. Subjects, instead, are of the discourse and involved in the construction of themselves as subjects in relation to the discourse (Dean, 1995). Perhaps most useful for the purposes of this paper is Foucault's (1980) writing on truth:

> Each society has its regime of truth, its "general politics" of truth: that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements, the means by which each is sanctioned; the techniques and procedures accorded value in the acquisition of truth; the status of those who are charged with saying what counts as true. (p. 131)

Foucault's articulation of truth provides a useful framework for thinking about the consequences of establishing a monolithic philosophy and practice of teaching within the Arizona and Texas systems.

During the final stage of analysis, I focused on problematising the "allure of alignment" (Savage & O'Connor, 2019, p. 21) by using Connolly's (2017) notion of pluralism to critique how TAP operates as an orthodox ontology, creating the rules, values and realities for teachers to practice, develop and exist within the system. Now, I move to the analysis, where I begin by describing the policy contexts of each state-based case.

The TAP System in Arizona and Texas

The Case of Arizona

Around the time that most states were working to secure Race to the Top grants, many local systems were also competing for federal funds that similarly encouraged accountability-based reform. In 2010, a team of 10 high-needs Arizona school districts along with Arizona State University and NIET applied for a federal Teacher Incentive Fund (TIF) grant. The project, called the *Arizona Ready-for-Rigor* project, acquired a $43.8 million grant to be used for their proposed five-year plan to implement the TAP System in the partner school districts and Arizona State University's teacher preparation program. Between 2010 and 2016, it is estimated that 100 administrators, 2,100 teachers, and 40,000 students were impacted by TAP annually in Arizona.[1] Furthermore, NIET, Arizona State University and 21 Arizona school districts partnered to build on the infrastructure established by the TIF project to apply for (and win) another federal grant (a SEED grant), called *Planting the SEED* ($12 million, 2013–2018). According to their application:

> The partnership between NIET and ASU allows for a comprehensive continuum of services from preservice to inservice that are based on strategies supported by strong evidence of effectiveness. Ultimately, the Project aims to implement practices in each of these areas that will be sustained by each partner district beyond the federal funding period, which will create a model that can be adopted by other agencies across the nation, including teachers' colleges and nonprofit organizations. (NIET, 2013a, p. 3)

It should be noted that Arizona State University's teacher preparation program is one of the largest colleges of education in the US, while TAP is one of the most prominent teacher evaluation systems. Next, I provide a brief description of the Texas case, which is strikingly similar to the Arizona scenario.

The Case of Texas

Texas has a long and contentious history with test-based accountability. In fact, No Child Left Behind has been largely credited to what US President George W. Bush called the "Texas Miracle" (or, more specifically, the Houston Independent School District "miracle;" see Leung, 2004). Texas teachers have thus faced high-stakes consequences for decades, including termination for

---

[1]   At the conclusion of the *Arizona Ready for Rigor* project, Arizona State University severed their ties with NIET and removed TAP from their program accordingly.

(at least in large part) failing to "add" enough "value" to their students' test scores on large-scale achievement tests (Amrein-Beardsley & Collins, 2012).[2]

In 2015, Texas Tech University, like Arizona State University, partnered with NIET and 5 school districts to apply for, and win, a federal SEED grant called the *TAP Connect National Pilot* project. The grant application described the project as follows:

> Technology-enabled, competency-based shaping will produce highly effective, "TAP-ready" new teachers, in-service TAP teachers with competency-based advanced certification in effective literacy and STEM instruction and TAP school leaders that effectively foster high-fidelity implementation of the TAP System. This technology-enabled, competency based approach will result in TAP schools that produce significantly higher student achievement than traditional TAP schools and local control schools. (NIET, 2013b, p. 2)

The application also promised that TAP Connect "would become an easily transportable, highly effective prototype for the entire country" (NIET, 2013b, p. 4). Indeed, at the conclusion of the TAP Connect project, Texas used the TAP System as its model for its state-wide teacher evaluation system (i.e., the Texas Teacher Evaluation and Support System), which was still in place at the time of writing this paper.

It should also be noted that the Dean of the College of Education at Texas Tech—who helped develop TAP Connect—had also worked at Arizona State University when it was collaborating with NIET for the Arizona *Ready-For-Rigor* and *Planting the SEED* projects. This is noted explicitly in the TAP Connect proposal and helps explain some of the obvious similarities between the two contexts. Like the TAP System in Arizona, the TAP System in Texas carefully aligns each domain of the Texas teacher's practice, professional status and development. In the following section, I bring these two cases together to (1) identify and map the instruments and practices within each of the teaching practice domains, and (2) illustrate how TAP operates as an orthodox ontology that shapes what is possible in terms of teachers and teaching within the Arizona and Texas systems.

---

[2]   In 2017, the Texas Education Agency required all districts to implement the *Texas Teacher Evaluation Support System*, or T-TESS. This is important for the current paper because much of T-TESS was modelled after the TAP System. Texas retained several TAP components as the foundation of what would eventually become the T-TESS system (T-TESS.org at https://www.teachfortexas.org). This comprehensive, state-level framework is not only an evaluation process but also an online platform (T-TESS.org), where data are stored, collected for analysis, and reported back to teachers and school leaders. T-TESS.org also provides a suite of resources related to the T-TESS evaluation process, including, but not limited to, training and calibration videos, observation rubrics, evaluation templates and procedure handbooks, many of which are TAP or NIET-created materials.

TAP Tools, Instruments and Procedures

The TAP System is an all-inclusive package that clearly details each stage of a teacher's development, appraisal and discipline. The system provides extensive handbooks, templates, videos, and training courses that leave no unanswered questions about what TAP Teachers should expect, or how TAP Evaluators should observe, coach, and assess teachers' (and pre-service teachers') performance. Perhaps the most prominent feature of the TAP System is the TAP rubric, which not only serves as the anchor of each domain, but also as the channel that connects each of the domains together. Table 1 illustrates the key instruments and procedures that define each of the teaching domains, which shows the significance of the rubric in particular. The table is followed by a description of each domain.

Table 1

TAP System Instruments and Procedures

Domain	Instruments and Procedures
Teacher preparation	• TAP rubric • Regular observations and TAP evaluations • TAP-aligned curriculum
Professional Development (PD)	• TAP rubric • Weekly sessions • Conducted by TAP evaluators
Evaluation	• TAP rubrics for observations of instruction and classroom environment • TAP rubrics for professional responsibilities outside of the classroom and self-assessment • Value-added measurement (based on student achievement test scores)
Conferencing, goal-setting and targeted coaching	• TAP rubric-based goal-setting (beginning of year) • Goal revisited throughout year (with TAP evaluator) • Scripted (TAP rubric-based) pre- and post-conferences (before and after formal observations) • Refinement and reinforcements (determined by rubric scores) are identified for intervention/coaching
Disciplinary Action	• Composite evaluation scores (rubric scores + valued-added scores) • Merit-based pay (based on rankings of evaluation scores) • Discipline plan or termination for consecutively inadequate evaluation scores • Promotion to leadership positions for high evaluation scores

Distributed leadership	• Teachers can be promoted to leadership positions (Mentor or Master teachers) • Mentors and Masters conduct evaluation, weekly cluster meetings, pre-/post-conferences, and targeted coaching with assigned teachers • Mentors and Masters also teach classes during the week
Video Observation (TAP Connect only)	• 'Mini Rigs', i.e., iPod Touch, tripod and 'high sensitivity microphone' (NIET, 2013b, p. 32). Mini rigs are to be used daily (or weekly at minimum) to record classroom practice and coursework • Videos are observed and evaluated by Regional Master Teachers and TTU faculty (using TAP rubric)

Teacher preparation

At Arizona State University and Texas Tech University, the teacher preparation programs are explicitly aligned to TAP, including the course content, the field teaching experiences, and the assessment of pre-service teachers. Pre-service teachers are expected to graduate "TAP-ready" (NIET, 2013b). The following excerpt from the *TAP Connect* application describes the degree to which TAP is embedded within the Texas Tech University teacher preparation program:

> The TAP instructional rubric for effective teaching is the framework that drives the *entire TTU teacher education program*. Each course is delivered live via interactive web interface to teacher candidates embedded in the partner TAP schools (in combination with an array of performance-based enrichment modules). Courses prepare teacher candidates to understand best-practices and to implement targeted indicators from the TAP rubric. The notion of "Apply & Evaluate" from the TAP Cluster Group process has been adopted in all TTU teacher education courses. This means that *every course requires teacher candidates to try out and video capture attempts to implement "best practices"* strategies in their TAP school classroom placements. These video captures are brought back to the teacher education class, shared virtually with and scored by peers on the TAP rubric. (NIET, 2013b, p. 25, emphasis added)

Professional development (PD)

Once teachers begin teaching in their schools (i.e., after graduation), they are required to attend weekly cluster, or professional development, meetings where they are taught field-tested lessons by Mentor or Master teachers. These lessons are aligned to the TAP rubric and are designed to improve TAP rubric scores. It should also be noted that pre-service teachers learn TAP lessons within their teacher preparation programs, where they are expected to implement such lessons in their field placement schools, record themselves teaching, and submit their recordings for peer-review. Lesson plans are also to be submitted for portfolio-based assessment, which is also aligned to the TAP rubric.

Evaluation

Teachers are observed formally up to four times a school year. TAP rubrics are used to assess the teachers' instruction, as well as the classroom environment during this time. Teachers then submit evidence of students' work and assessment so that rubrics can be used to assess the teachers' ability to assess student performance. Teachers are also required to use rubrics to self-assess their own performance. The self-assessment is compared to the evaluator's assessment, which is then used to assess whether the teacher is proficient at self-assessing. Teachers are also evaluated by value-added measurement, which is the measurement of growth that students make on standardised achievement tests, over time, and then attributed to their teachers. Rubric and value-added measures are combined for final evaluation scores.

Conferencing, goal-setting and targeted coaching

Teachers are required to identify a TAP-based goal at the beginning of the year, which is regularly revisited during one-on-one conferences. Additionally, after formal observations, the evaluator identifies an area of refinement (to improve) and an area of reinforcement (to continue) for the teacher. This is also based on the TAP rubric and the TAP-based assessment of the teacher's performance (i.e., TAP indicators that received low or high scores). The teacher might receive targeted coaching (or particular strategies) to improve their area of refinement for their next observation.

Disciplinary action

At the end of the school year, the teacher's composite evaluation score is used for ranking the teachers against their colleagues and then distributing performance-based pay accordingly. Teachers can also be placed on discipline plans or terminated for failure to improve evaluation scores (this varies across the two states).

Distributed leadership

Teachers can be rewarded for good TAP performance by being promoted into leadership positions. Mentor and Master teachers are peer evaluators who are responsible for teaching, but also observing, evaluating, and coaching their assigned teachers. Mentor and Master teachers receive small stipends for their additional responsibilities, and they serve on the leadership team of their school. They teach the weekly cluster meetings, and they conduct evaluations and pre- and post-conferences with teachers.

As is made clear in these descriptions, there is a lot of overlap and redundancy within each of these domains. In doing so, there is never a question about expectations, including how teachers should conduct themselves, as well as what will happen if they fail to do so adequately (e.g., missed opportunities

for extra pay, being placed on discipline plans, failure to pass university courses, etc.). Furthermore, such clear expectations and consequences also mean that there is little room for teachers to deviate from the course. On one hand, this supports consistency across the system; however, on the other hand, this minimises opportunities for plurality or change. Thus, I argue that the way the TAP System operates within these systems creates an *orthodox ontology*, which I develop in the following section.

The Orthodox Ontology of TAP Teachers and Teaching

First, let me define what I mean by orthodox ontology. Merriam-Webster Dictionary defines orthodox as "conforming to established doctrine especially in religion." As mentioned earlier, my argument is situated within a Foucauldian perspective that defines discourse as frameworks of thought that make and define possibilities. In other words, "discourse" is that which constitutes the knowable and the imaginable, operating as a "regime of truth" (Foucault, 1980) rather than truth itself (McWilliam & Jones, 2005). Here, discourse is the language, practices, and fields of knowledge that operate to construct our reality, and which are constantly being negotiated and renegotiated in any given moment and space. Accordingly, we must work to understand how language, over time, shapes reality and constitutes particular ways of knowing, doing, and being. This pertains to the current analysis in that the TAP System, and its ubiquity throughout all stages and dimensions of the pre-service and practicing teachers' experiences, creates an onto-epistemic framework (cf. Holloway, 2019) that constitutes what is valued, knowable, and imaginable for TAP teachers. As the transitions between the various domains (e.g., training, evaluation, professional, etc), are made evermore seamless—as is a main objective of the TAP grants—then the exposure the teachers have to other discourses, knowledges, practices, and so forth, about teaching is increasingly reduced.

Perhaps a bit provocative, but like a religious doctrine, TAP is treated as the fundamental creed that underpins and defines teaching practice for these grant participants. By design, the teachers are first prepared in universities that use the TAP rubric as the core of their teaching framework, and then they enter schools that also use the TAP rubric as the core of their professional development and evaluation system (NIET, n.d., p. 1). It is arguable, then, that these teachers come to experience TAP as the sole definition of good teaching practice. In the name of alignment and standardization, this system creates a closed circuit that makes imagining or practicing other forms of teaching difficult, if not impossible. As argued by NIET:

> States and districts must create their own plans for ensuring that teachers are
> prepared to teach to the higher standards. This requires states to rethink
> their professional development, support and evaluation systems to adequately
> train and support teachers. Based on more than a decade of experience in
> developing and implementing systems of teacher support and evaluation,
> NIET has found that aligning these systems with instruction can put schools
> on the right path towards increased teacher effectiveness and student
> achievement growth. (NIET, 2013, n.p.)

This is not to say that alignment, in general, is necessarily bad, but strict
alignment and efforts to create a smooth pipeline for teachers to transition
from the university to schools does present possible problems in terms of
exposure to plural philosophies about teaching and different ways of
responding to a variety of student needs.

Indeed, TAP's ubiquitous presence in these interrelated domains helps
build system alignment, but it also means that teachers are increasingly limited
in their exposure to alternative philosophies, practices, and measurements
of teaching. Each stage of their professional identity is calibrated to TAP,
which is characterised by high-stakes accountability (e.g., value-added
evaluation, merit pay) and standardisation (e.g., rubric-based observation and
professional development).

When Alignment Stifles Plurality

Thousands of Arizona and Texas teachers have passed through TAP-based
preparation programs, only to enter TAP-based schools. Indeed, it is a key
priority of these grants that teachers leave their universities ready for TAP,[3]
which compliments broader standards-based reform efforts around consistency
and alignment (see Lee, 2019; US Department of Education, 2016). However,
it is equally important to consider the potential dangers associated with strictly
defining (and confining) that which teachers experience in their preparation
programs.

---

[3]   It should also be noted that the main US accreditation program—Council for the
       Accreditation of Educator Preparation (see caepnet.org)—also requires that teachers
       leave their programs fully prepared to enter the accountability environments of K-12
       schools. This is measured and assessed via survey instruments filled out by graduate
       teachers and their principals after they leave their programs. Such data are used to
       assess the quality and integrity of the university-based teacher preparation programs.

Strict alignment between teacher preparation, evaluation, and professional development creates a narrowly defined view of teaching that makes adaptability and professional discretion difficult, if not impossible, for teachers to practice. Importantly, this is not to say that TAP in and of itself is based on a narrow view of teaching. Rather, as TAP presides as a multi-domain system, the boundaries between standard (i.e., TAP) and alternative practice grows increasingly rigid, which impedes teachers' capacities to be critically-responsive to the changing needs of their students (cf. Connolly, 2017). In other words, the constant reinforcement that TAP is the only way to teach means that teachers are minimally equipped to respond to changing circumstances. This process can be viewed as a positive move towards building system coherence, making it easier for teachers to know what good teaching entails and the know-how to become a good teacher (at least in terms of how they will be assessed on their evaluations). In fact, many researchers—particularly from the field of educational measurement —posit that alignment between standards and assessment are necessary for achieving content-related validity (see Martone & Sireci, 2009). In a practical (and admittedly oversimplified) sense, their argument rests on the assumption that clear evaluative criteria are necessary for measuring the degree to which someone adheres to, or performs against, the said criteria. Theoretically, this can lead to improvement in teaching, as teachers have a clear understanding of what is expected of them in terms of their performance (this is one of the main objectives of the TAP System; NIET, n.d.). If we think of this assumption through the lens of Foucault's "regime of truth," though, we can see that aligning evaluation criteria (i.e., "normalising judgment") with the practice of evaluation (i.e., "examination") constructs improvement in relation to the evaluation regime. In other words, improvement is not a real or stable construct, but a product of the particular conditions at play. Furthermore, we must question what this means for facilitating diversity of thought and behaviour, which is often a valued principle of western liberal democracies. In some ways, TAP smooths out the wrinkles of an otherwise messy practice like teaching – it defines clear expectations, gets everyone on the same page, provides benchmarks and well-defined targets for teachers to achieve. Yet, if we take an alternative perspective, we might ask how some degree of disorder, misalignment and messiness can be productive for democratic institutions, like schools. I find Savage and O'Connor's (2019) argument on this matter particularly compelling:

> …we see alignment thinking as tending towards erasing from view any forms of disorder that might prove to be incapable of being aligned. This is done through privileging standardisation over difference, commonality over diversity, and connections over disjuncture. Arguments for policy alignment

> thus exemplify core elements of Scott's articulation of 'high modernist' attempts to impose rational and technical order onto social worlds that are inherently messy, complex and resistant to ordering. (Scott, 1998; Savage & O'Connor, 2019, p. 8)

Teaching is a messy and multifaceted process that requires complex approaches and philosophies, especially when we consider the current climate within which teachers are asked to teach. From globalisation and automation, to mass immigration, climate change, or the rise of right-wing extremism, teachers sit at the front lines when it comes to the social, political, and environmental challenges that societies face (Hursh et al., 2015; Komatsu et al., 2019; Subedi & Daza, 2008). As such, efforts to simplify and make rigid what counts as good teaching is inherently at odds with what we know about the practice, as well as what we know about the dynamism of current conditions. Thus, exercising a bit of scepticism when it comes to a monolithic system of authority is vital for creating the cracks through which new possibilities and imaginaries of teaching might emerge during these turbulent times. The TAP System, instead, creates a smooth pipeline that purposefully seals any such cracks, removing the possibility that teachers might become any other than a TAP Teacher. This may or may not be what is necessary for the unprecedented obstacles that teachers are inevitably to encounter, but what it does do is reduce the capacity for flexibility and adaptability, should the changing conditions require it.

These concerns also resonate with what scholars of New Public Management [NPM] (see Anderson & Herr, 2015; Zeichner, 2010) and the de-professionalising (or re-professionalising) of teachers (Brass & Holloway, 2019; Sachs, 2001; Taubman, 2009) have been raising for some time now. These scholars point to how the teaching profession has shifted from an occupational to organisational form of professionalism, where metrics, standards, and hierarchical control of teachers' work has drastically curtailed (or thoroughly re-constituted) teachers' authority and autonomy in the classroom. The TAP System is an exemplar case of this new managerialism.

I should note that I am mindful of how my argument might sound – that naïve teachers are being brainwashed to follow TAP. I want to be clear that this is not what I am saying. What I am saying is that, for many of these teachers, they receive the TAP rubric from the first day they begin training to become a teacher; their graduation and certification depends on adherence to the rubric; once they begin teaching, their professional goals and weekly professional development programs are based on the rubric; their classroom practice is evaluated by how well they adhere to and embody the rubric; and their promotion, retention and pay is calculated on their rubric-based appraisal. This process does not brainwash teachers, but it does impose

a single belief system that must be followed – otherwise teachers risk symbolic and material (high-stake) consequences (see also Holloway, 2019). I am reminded of Courtney and Gunter's (2015) provocative work on visioning and vision work (i.e., the practice of developing and implementing a vision for an institution). They borrow from Arendt's (1958; 2009) historical and political view of totalitarianism to critique the dangers of enforcing a singular vision for a school, especially when the conditions are such that teachers must comply. Like the visioning practice and purpose described by the school leaders in their study, the TAP System (1) operates as the presiding ideology, (2) uses terror to discipline teachers through sanctions, rewards, and threats of termination, (3) weakens human bonds by supplanting interpersonal relationships with performance regimes, data dashboards, and standards (see also Garver, 2019, on "evaluative relationships"), and (4) incorporates TAP-mediated bureaucratic controls within each stage of teacher certification, development, evaluation, and discipline. In other words, there is no escape or alternative vision for TAP teachers to explore or enact. Whether the teachers, or even the professors who teach in the TAP-based preparation programs, believe in the benefits of the TAP System is somewhat irrelevant. The way that TAP operates renders alternative perspectives insignificant insofar as any deviance from TAP is met with tangibly severe consequences. This begs the question, how does TAP not function as a totalising regime?

Conclusion

While there is a robust literature on the various debates about, and consequences resulting from, new accountability systems and mechanisms (see, for example, American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; American Statistical Association, 2014; Amrein-Beardsley & Holloway, 2019), my argument departs from this literature to problematise the discursive and material conditions that have produced a sort of TAP-based orthodoxy (cf. Bacchi, 2012). If we view this arrangement through a Foucauldian lens, we can begin to see the ways that multi-domain accountability systems (like TAP) that prioritise alignment and high-stakes accountability operate as a "regime of truth" (Foucault, 1980) that profoundly limits teachers' capacities to think outside of the TAP box. As such, I have attempted to disrupt the common assumptions associated with goals of strict alignment by highlighting how an enclosed system prevents teachers from accessing new ways of thinking about themselves as teachers or their teaching practice.

To this point, I suggest finding ways of acknowledging and embracing the complexity that is inherent in teaching, and especially when teaching diverse student populations. I urge that we consider seeing "productive disorder" (Savage & O'Connor, 2019, p. 8) as a means for building conditions that enable teachers and schools to be critically responsive to the changing needs of their students (Connolly, 2005), as well as for creating space for imagining new ways of teaching during contemporary (and possibly tumultuous) times. This is particularly true in places like Arizona and Texas, where student demographics and needs are rapidly changing (Collier & Ura, 2015; Creno, 2016).

Furthermore, it is worth viewing the motivation to align systems and clean out any messiness in terms of what this means for dissension (Sachs, 2001). If teachers are forced to view TAP as the authoritative ideology on teaching, then possibilities for being other, practicing alternatively, or advocating for change become quite difficult, as the system works from every angle to symbolically and materially structure the conditions of possibility, or the ontological space, for teacher behaviour and being to become and exist. I argue that the TAP System thus creates a fixed ontological space that sets the purpose, identity, and function of the teacher subject, but also erases alternative channels through which other types of teachers might become. While producing TAP-ready teachers might be lauded as an ideal objective of the grants, it is critical that we challenge the assumptions associated with these types of endeavours. This is not a critique of the quality of the universities involved in this study, nor of the TAP or TAP instruments, practices, or philosophies upon which TAP is founded. What I am questioning is whether creating a closed space within which teachers are made, disciplined and operate is the best way to promote diversity of thought and practice that is necessary for accepting the complexity of the current time. There is careful consideration that needs to be done in terms of prioritising alignment at the expense of plurality – for plurality is vital for establishing critically-responsive institutions (Connolly, 2005), especially in changing times and in dealing with diverse populations.

Acknowledgement

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

American Statistical Association. (2014). *ASA statement on using value-added models for educational assessment.* https://www.amstat.org/policy/pdfs/ASAVAMStatement.pdf

Amrein-Beardsley, A., & Collins, C. (2012). The SAS education value-added assessment system (SAS® EVAAS®) in the Houston Independent School District (HISD): Intended and unintended consequences. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, *20*(1), 1–28. https://doi.org/10.14507/epaa.v20n12.2012

Anderson, G., & Herr, K. (2015). New public management and the new professionalism in education: Framing the issue. *Education Policy Analysis Archives*, *23*(84), 1–5. https://doi.org/10.14507/epaa.v23.2222

Arendt, H. (1958*). The human conditions.* Doubleday.

Arendt, H. (2009). *The origins of totalitarianism (2nd ed.).* Benediction Books.

Bacchi, C. (2000). Policy as discourse: What does it mean? Where does it get us? *Discourse: Studies in the Cultural Politics of Education*, *21*(1), 45–57. https://doi.org/10.1080/01596300050005493

Bacchi, C. (2012). Why study problematizations? Making politics visible. *Open Journal of Political Science*, *2*(1), 1–8. http://dx.doi.org/10.4236/ojps.2012.21001

Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, *18*(2), 215–228. https://doi.org/10.1080/0268093022000043065

Ball, S. J. (2015). Education, governance and the tyranny of numbers. *Journal of Education Policy, 30*(3), 299–301. https://doi.org/10.1080/02680939.2015.1013271

Bradbury, A. (2019). Datafied at four: The role of data in the 'schoolification' of early childhood education in England. *Learning, Media and Technology*, *44*(1), 7–21. https://doi.org/10.1080/17439884.2018.1511577

Bradbury, A., & Roberts-Holmes, G. (2017). *The datafication of primary and early years education: Playing with numbers.* Routledge.

Collier, K., & Ura, A. (2015). *Texas public schools are poorer, more diverse.* Texas Tribune. https://www.texastribune.org/2015/12/09/new-schools-explorer-shows-changing-face-schools/

Collins, C., & Amrein-Beardsley, A. (2014). Putting growth and value-added models on the map: A national overview. *Teachers College Record*, *116*(1), 1–32.

Connolly, W. E. (2013) *The fragility of things: Self-organizing processes, neoliberal fantasies and democratic activism.* Duke University Press.

Connolly, W. E. (2005). *Pluralism.* Duke University Press.

Connolly, W. E. (2017). The challenge to pluralist theory. In W. E. Connolly (Ed.), *Pluralism in political analysis* (pp. 3–34). Routledge.

Courtney, S. J., & Gunter, H. M. (2015). Get off my bus! School leaders, vision work and the elimination of teachers. *International Journal of Leadership in Education*, *18*(4), 395–417. https://doi.org/10.1080/13603124.2014.992476

Creno, C. (2016). *Arizona's changing demographics and the academic divide.* Helios education foundation. https://www.helios.org/blog/part-i-arizona%E2%80%99s-changing-demographics-and-the-academic-divide

Daley, G., & Kim, L. (2010). *A teacher evaluation system that works*. National Institute for Excellence in Teaching. http://www.niet.org/assets/Publications/teacher-evaluation-system-that-works-working-paper.pdf

Daley, G., & Kim, L. (2012). *TAP research summary*. National Institute for Excellence in Teaching.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. ASCD.

Davies, B., & Bansel, P. (2010). Governmentality and academic work: Shaping the hearts and minds of academic workers. *Journal of Curriculum Theorizing*, *26*(3), 5–20.

Dean, M. (1995). Governing the unemployed self in an active society. *International Journal of Human Resource Management, 24*(4), 559–583. https://doi.org/10.1080/03085149500000025

Foucault, M. (1977). *Discipline and punish: The birth of the prison*. Vintage.

Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings 1972–1977*. Pantheon.

Foucault, M. (2000). Truth and juridical forms. In J. D. Faubion (Ed.), *Michel Foucault: Power* (pp. 1–89). The New Press.

Foucault, M. (2002). *The order of things: An archaeology of the human sciences*. Routledge.

Garver, R. (2019). Evaluative relationships: Teacher accountability and professional culture. *Journal of Education Policy*, 1–25. https://doi.org/10.1080/02680939.2019.1566972

Hardy, I. (2019). The quandary of quantification: Data, numbers and teachers' learning. *Journal of Education Policy*, 1–20. https://doi.org/10.1080/02680939.2019.1672211

Hewitt, K. K., & Amrein-Beardsley, A. (Eds.). (2016). *Student growth measures in policy and practice: Intended and unintended consequences of high-stakes teacher evaluations*. Springer.

Holloway, J. (2019). Teacher evaluation as an onto-epistemic framework. *British Journal of Sociology of Education, 40*(2), 174–189.

Holloway, J., & Brass, J. (2018). Making accountable teachers: The terrors and pleasures of performativity. *Journal of Education Policy, 33*(3), 361–382.

Hursh, D., Henderson, J., & Greenwood, D. (2015). Environmental education in a neoliberal climate. *Environmental Education Research*, *21*(3), 299–318. https://doi.org/10.1080/13504622.2015.1018141

Komatsu, H., Rappleye, J., & Silova, I. (2019). Culture and the independent self: Obstacles to environmental sustainability? *Anthropocene*, *26*, 100198. https://doi.org/10.1016/j.ancene.2019.100198

Lee, O. (2019). Aligning English language proficiency standards with content standards: Shared opportunity and responsibility across English learner education and content areas. *Educational Researcher*, *48*(8), 534–542. https://doi.org/10.3102/0013189X19872497

Lingard, B., & Rizvi, F. (2010). *Globalizing education policy*. Routledge.

Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, *79*(4), 1332–1361. https://doi.org/10.3102/0034654309341375

McWilliam, E., & Jones, A. (2005). An unprotected species? On teachers as risky subjects. *British Educational Research Journal*, *31*(1), 109–120. https://doi.org/10.1080/0141192052000310056

NIET. (2011). *Reliability in Classroom Observations*. Presentation by the National Institute for Excellence in Teaching, October 2011. https://cepr.harvard.edu/files/cepr/files/ncte-webinar-niet-powerpoint.pdf

NIET. (2013a). *Planting the SEED*. Grant application for Supporting Effective Educator Development Grant Program, Santa Monica CA. https://www2.ed.gov/programs/edseed/2013/nietasu.pdf

NIET. (2013b). *TAP Connect National Pilot*. Grant application for Supporting Effective Educator Development Grant Program, Santa Monica CA. https://www2.ed.gov/programs/edseed/2013/nietttu.pdf

NIET. (2013c). *TAP system prepares teachers for common core*. National Institute for Excellence in Teaching.

NIET. (n.d.) *Rubric and observation systems*. National Institute for Excellence in Teaching. https://www.niet.org/our-work/our-services/show/rubric-and-observation-systems

Perryman, J. (2009). Inspection and the fabrication of professional and performative processes. *Journal of Education Policy*, *24*(5) 611–631. https://doi.org/10.1080/02680930903125129

Sachs, J. (2001). Teacher professional identity: Competing discourses, competing outcomes. *Journal of Education Policy*, *16*(2), 149–161. https://doi.org/10.1080/02680930116819

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.

Savage, G. C., & Lewis, S. (2018). The phantom national? Assembling national teaching standards in Australia's federal system. *Journal of Education Policy*, *33*(1), 118–142. https://doi.org/10.1080/02680939.2017.1325518

Savage, G. C., & O'Connor, K. (2019). What's the problem with 'policy alignment'? The complexities of national reform in Australia's federal system. *Journal of Education Policy*, *34*(6), 812–835. https://doi.org/10.1080/02680939.2018.1545050

Scott, J. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.

Sloat, E., Amrein-Beardsley, A., & Sabo, K. E. (2017). Examining the factor structure underlying the TAP system for teacher and student advancement. *AERA Open*, *3*(4), 1–18. https://doi.org/10.1177/2332858417735526

Subedi, B., & Daza, S. L. (2008). The possibilities of postcolonial praxis in education. *Race Ethnicity and Education*, *11*(1), 1–10. https://doi.org/10.1080/13613320701845731

Taubman, P. M. (2010). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education*. Routledge.

The Danielson Group. (n.d.). *The Danielson Framework website*. https://danielsongroup.org/

US Department of Education. (2016). *Every student succeeds act*. https://www.ed.gov/essa?src=rn

Zeichner, K. (2010). Competition, economic rationalization, increased surveillance, and attacks on diversity: Neo-liberalism and the transformation of teacher education in the U.S. *Teaching and Teacher Education*, *26*(8), 1544–1552. https://doi.org/10.1016/j.tate.2010.06.004

Corresponding author
Jessica Holloway
Research for Education Impact, Faculty of Arts and Education, Deakin University, Geelong, Australia
E-mail: jessica.holloway@deakin.edu.au