When Logics Collide: Implementing Technology-Enabled Personalization in the Age of Accountability

Iris Daruwala[1], Shani Bretas[1], and Douglas D. Ready[1]

In this article, we situate the relatively new wave of technology-enabled personalized learning platforms within the broader policy context of institutional accountability. Although many education technology innovations were developed in response to heightened accountability pressures, technology-enabled personalization introduces new institutional rules, cultures, and norms that may be at odds with the goals of accountability. Using the adoption of a personalized, blended learning mathematics program at five U.S. schools as a case study, we describe how teachers, school leaders, and program staff navigated institutional pressures to improve state grade-level standardized test scores while implementing tasks and technologies designed to personalize student learning.

Keywords: accountability; educational policy; organizational theory; personalized learning

An emerging approach to school improvement has been steering towards a collision with the institutional environment of accountability in recent years. On one hand, accountability systems have forced most public schools to accept an institutional logic rooted in grade-level assessments and proficiency benchmarks. On the other hand, districts and schools are increasingly adopting models that personalize instruction, focusing on relative measures of cognitive development rather than absolute levels of achievement. There is a clear discrepancy between the homogeneous expectations within traditional accountability environments and the heterogeneous interpretations of student growth that undergird personalized approaches. These competing notions leave school personnel to grapple with a fundamental question: How do we implement personalized instruction in an environment that demands grade-level outcomes?

In this article, we identify technology-enabled personalization as a nascent organizational field—or a distinct community of organizations that partake in a shared meaning system (Scott, 1995)—with an institutional logic sometimes at odds with the prevailing logic of accountability. Using evidence from the implementation of one technology-enabled personalization model, we argue that these tensions have shaped individual sensemaking, routines, and organizational practices in ways that undermine the coherence of both accountability and personalization. We examine how teachers, school leaders, and program designers attempted to reconcile the contradictory expectations placed upon them during implementation. The purpose of our study is not to simply identify these conflicting demands and their impact on stakeholders, but to examine how their discrepant logics contribute to the cyclical, and even counterproductive, nature of education reform.

Background

The Logic of Accountability

Scholars have long sought to understand why organizations behave the way that they do. Friedland and Alford (1991) first introduced the concept of "institutional logics" to describe how different institutions contribute to rationalizations that organize behavior. This behavior in turn creates routines and rituals that ultimately reproduce widespread belief in such institutions. These logics are not simply imposed, but are coconstructed, strengthened, and preserved through individual, organizational, and institutional interactions (Haveman & Gualtieri, 2017). As many scholars have documented, the logic of accountability can be broadly characterized as a shift towards a more technical educational environment that holds educators and leaders responsible for student outcomes (Hallett, 2010; Lowenhaupt et al.,

[1]Teachers College, Columbia University, New York, NY

Educational Researcher, Vol. 50 No. 3, pp. 157–164
DOI: 10.3102/0013189X20960674

2016; Meyer & Rowan, 2006; Rowan, 2006). Numerous researchers have explored the routines and rituals—or "myths and ceremonies" (Meyer & Rowan, 1977)—that schools have adopted in order to conform to this logic of accountability. Some have observed a shift away from the traditional loose coupling found in education—when schools enjoyed considerable local autonomy and faced few statutory demands (Weick, 1976)—to a tighter coupling of instructional tasks and technologies associated with institutional pressures and sanctions for noncompliance (Meyer & Rowan, 2006). Regardless of how schools adapt to external accountability pressures, however, fundamental to the logic of accountability is the premise that legitimacy is determined through commensurate outputs: grade-level proficiency benchmarks.

Despite these institutional expectations of commensuration, teachers typically confront tremendous variability in student skills within the classroom. For example, estimates suggest that almost 62% of the variability in fifth-grade mathematics ability exists within classrooms, with only 18% of the variance lying between teachers in the same school, and 20% between schools (Martinez et al., 2009). Recent advances in computing technology have increased interest in personalized instructional approaches as a solution to managing heterogeneity while simultaneously adhering to broader accountability expectations of grade-level proficiency (Bill & Melinda Gates Foundation, 2014; Wolf, 2010).

Although many proponents argue that using technology to personalize the classroom for individual learners is an effective means to achieve the goals of grade-level accountability (Bill & Melinda Gates Foundation, 2014), elements of these models may also be fundamentally at odds with the notion of accountability. Namely, technology-enabled personalization emphasizes tailoring instruction to individual learning needs and measuring growth against individual baselines rather than uniform benchmarks. Although the ultimate aim may be grade-level achievement, students embark on individualized paths to accomplish that goal, resulting in timelines that often fail to coincide with the typical yearly testing cycle. Thus, despite the shared goals of improving instruction and student performance, technology-enabled personalization diverges greatly from the logic of accountability in the means through which it proposes to accomplish such goals.

Sensemaking Around the Introduction of New Tasks and Technologies

When faced with multiple and potentially contradictory tasks, technologies, and expectations, increased ambiguity can compromise how individuals make sense of their work. We refer to the process of "sensemaking" as a combination of individual cognition and social interactions built on the connection between "frames," or historical reference points; and "cues," or present experiences (Weick, 1995). In complex environments, individuals rely on subconscious schemas—classification frameworks that organize thoughts and interpretations—to simplify and minimize cognitive effort (Lizardo et al., 2016; Rosch, 1978; Strauss & Quinn, 1997). These latent schemas shape how and when individuals recall frames and respond to cues.

As schools adopt multiple policies, often with different goals and processes, teachers and school leaders are likely to rely on routine frames and cues to process increasingly diverse information (Spillane et al., 2002; Weick, 1996). However, the introduction of conflicting beliefs, norms, and practices can weaken the ability of individuals to simplify and classify information into schematic constructs, disrupting the traditional frames and cues individuals would normally call on to make sense of an event and respond appropriately (DiMaggio & Powell, 1991; Louis, 1980; Vaughan, 1996; Weick, 1995, 1996). For this reason, examining how school stakeholders make sense of their roles, responsibilities, organizational technologies, and tasks during periods of uncertainty is critical to understanding how cognitive and cultural coherence influence the implementation of both accountability and personalization policies. With sensemaking compromised, there is room for confusion or chaos that inhibits both policies from playing out as intended.

Research Focus

The idea that multiple, conflicting institutional logics exist within an organization's environment is not a particularly novel concept (e.g., see Glazer et al., 2018; Hallett, 2010), but surprisingly few researchers have explored how personalized learning approaches intersect with existing institutional environments. Our study aims to first identify technology-enabled personalization as an organizational field with its own institutional logic. Further, we explore the ways in which discrepancies between the institutional logics associated with accountability and technology-enabled personalization contribute to disruptions in sensemaking, organizational processes, and structures, using evidence from the implementation of one technology-enabled personalization program. Specifically, we address the following research questions:

1. To what extent can technology-enabled personalization be considered an organizational field with its own logic?
2. How do teachers, school leaders, program staff, and the creators of one technology-enabled personalization model navigate the implementation of personalized tasks and technologies within an institutional environment of accountability?
3. To what extent does technology-enabled personalization conflict with the logic of accountability?

Data and Methods

As a vehicle to explore the potential tensions between traditional notions of school accountability and more recent technology-enabled personalized learning platforms, we use the recent implementation of an anonymous technology-enabled personalized program, which we will refer to as TEPP. These analyses employed qualitative data from five K–8 schools implementing TEPP in their upper grades over 3 academic years, beginning in fall 2015. These schools were located in a midsized, high-minority-enrollment, high-poverty school district in the northeastern United States. The vast majority of students were either Black or Latino, and most were eligible for free or reduced-price lunch.

Every spring of the 3-year implementation, we conducted qualitative field research in each of the five participating schools. In our visits to each school, we interviewed school leaders, lead math teachers, TEPP teachers, and TEPP central program staff members. In Years 2 and 3, we also conducted separate focus groups with students at each of the five participating schools. In total, we conducted 121 interviews and 27 focus groups over the course of 3 years. Interviews and focus groups were semistructured; lasted approximately 30 to 45 minutes each; and addressed impressions of the model, how it was working for adults, and how it was working for students. Importantly, we managed to interview nearly all participating math teachers across the five schools once each year over the course of the 3-year implementation. We coded each interview separately using the Atlas.ti software, analyzed response patterns across the interviews, and wrote detailed memos summarizing our findings from each year. After our initial round of coding, we identified themes that emerged from the data and created "etic codes"—or codes that reflect our own analysis and thinking about participants' responses (Maxwell, 2012)—to capture interesting organizational patterns in a second round of analysis.

Findings

Context: The Emerging Field of Technology-Enabled Personalization

In recent years, technology-enabled personalization programs have been gaining traction in part due to heavy investment from large foundations such as the Bill & Melinda Gates Foundation, the Chan Zuckerberg Initiative, the Michael & Susan Dell Foundation, and Emerson Collective (Gross & DeArmond, 2018; Herold, 2016). Since 2012, there has been a marked increase in the number of districts requesting bids for personalized learning services (Molnar & Herold, 2018). Further, 15 states have taken legislative or regulatory steps to support personalized learning in some capacity (Burnette, 2017). As a result, a growing number of schools and districts across the country have adopted variations of such approaches (Hyslop & Mead, 2015), despite the dearth of rigorous research on the effectiveness of different models.

One challenge facing our effort to define technology-enabled personalization as a unique organizational field is that neither research nor practice has coalesced around a clear, shared definition. Further complicating matters is the recent surge in the number and types of personalized platforms. This profusion of new models has clouded public understanding of the differences and commonalities across various approaches. For example, although the terms "blended learning" and "personalized learning" are often used interchangeably, they actually represent quite distinct (yet frequently overlapping) constructs; a school may be blended, using both technology-enabled and live teacher instruction, without being personalized, or personalized without being blended, and technology may or may not play a role. In addition, many have used the term "competency-based learning" synonymously with both personalized learning and blended learning, although definitions typically include competency-based advancement as only one element of a broader conceptualization of personalization. Moreover, the technology-enabled personalized program we profile in this study itself did not incorporate every element that proponents of personalization would deem essential. For example, students were not afforded agency in determining either what or how they would learn each day and because schools employed this model as their core mathematics curriculum, there was no flexibility in when or for how long students learned mathematics each day. All students, regardless of interest or ability, encountered the platform whenever their daily class schedules dictated they should do so.

Despite these disagreements over "personalization" definitions and terms, we argue that a strong underlying logic unites the tasks and activities that fall under the "technology-enabled personalization" umbrella. In particular, technology-enabled personalization approaches entail individualized learning plans based on student-level data, recognize progress that is based on demonstrated knowledge rather than seat time, measure growth against individual student baselines, and employ multiple and flexible pedagogical and learning environments (Bill & Melinda Gates Foundation, 2014). These shared practices form part of a common meaning system that distinguishes technology-enabled personalization as its own organizational field, despite variability across models. Just as the variability in the particulars of state and district accountability systems do not weaken the core tenets of the logic of accountability, variations in technology-enabled personalization approaches do not lessen the sense among designers and practitioners that they are engaged in a common endeavor.

Organizational fields, through processes of individual, organizational, and institutional interactions rooted in a shared meaning system, often form and enact their own logics. This was evident at our five case study schools, where the TEPP program required significant changes to the traditional classroom environment. TEPP physically reorganized the learning environment into one large room containing multiple teachers and up to 100 students. Upon entering the room, students met in small advisory sections; opened personal laptop computers; logged into the TEPP online portal; and consulted their personal "playlists," which told them what content they would be learning that day and how they would be learning it. This blended program incorporated a variety of instructional modalities, including computer-based and pencil-and-paper independent work, peer-to-peer and small group learning, and traditional teacher-led instruction. Large TV screens directed students to designated areas for their first assigned modality and students again consulted these screens halfway through the TEPP period for their second assignment. At the end of each day, students returned to their advisories for a short, multiple choice "exit slip" to determine their mastery of the day's content.

Intended to differentiate instruction, the TEPP algorithm used students' exit slip results, combined with their baseline and ongoing benchmark assessments, to determine their skill and modality assignments for the next day. The algorithm intentionally assigned each student to a balance of modalities. At any given time, students in TEPP classrooms could have been

working independently, engaged in peer-to-peer learning, or participating in teacher-led instruction. Importantly, these students might have been working on mathematics content multiple grade levels above or below their assigned grade. As a result, students in the same room and even the same modality could be experiencing vastly different content on any given day. This physical reorganization of the classroom, reconfiguration of teachers' roles and responsibilities, and algorithm-driven personalization of student activities based on individual test scores form the core logic rationalizing teacher and student behavior under TEPP.

Collision: Tensions Between TEPP and Accountability

In the decade prior to TEPP implementation, the school district had conformed to the expectations of the accountability environment through the adoption of two practices: state-mandated administration of grade-level assessments and a statewide teacher evaluation system based on value-added measures of student performance and classroom observations. Prior to TEPP implementation, the five participating schools had struggled on the state assessments, with only about one-quarter of students meeting or exceeding the state grade-level standards in mathematics. Under the high-stakes teacher evaluation system, teachers were held accountable for the achievement of their students. Poor student performance could seriously impact teacher job stability, leading to possible contract termination. Given this high-stakes environment, teachers and leaders were initially excited about the promise of TEPP. They imagined that the program would build students' foundational skills and ultimately increase student achievement.

However, as TEPP implementation progressed, it became increasingly evident that the personalized nature of the program had its own logic—or reinforcing beliefs, values, tasks, and technologies—that conflicted with the accountability environment in which the schools and district were enmeshed. Principals, teachers, and even some students were forced to find ways to reconcile their general support for a personalized learning philosophy with the realities of the broader education policy environment. In the sections that follow, we describe these tensions through the eyes of teachers and school leaders, recounting their attempts to reconcile what many viewed as incompatible approaches to both pedagogy and policy.

Assessment. The most pervasive manifestation of the accountability logic in the district were state-mandated assessments intended to measure student mastery of grade-level content. The intention of these assessments was in direct conflict with the defining characteristic of the TEPP platform: students missing foundational math skills were provided the time and space to work through content *below* their current grade level if needed, thereby reducing the time spent on grade-level content. This had two serious implications for TEPP schools. First, regardless of instructional approach, grade-level assessments are often unable to capture academic growth made by students at the high and low tails of the achievement distribution, masking improvement. This fact was particularly salient to the district, which enrolled substantial proportions of very low-achieving students. Second, even with assessments that are able to capture academic growth across a range of initial student abilities, the differentiation inherent in the model likely exposed TEPP students to quite different mathematics content during the academic year than they would have experienced in a traditional mathematics classroom. Specifically, low-performing students typically received increased amounts of grade-level content under traditional mathematics curricula compared to what they received in TEPP. Greater exposure to grade-level content may provide a short-term positive impact on both grade-level assessments and those measuring growth. Given this, TEPP schools risked appearing less successful by grade-level standards, and therefore less legitimate than schools with traditional mathematics curricula under the logic of accountability.

Nearly all principals and teachers in our study expressed concerns about their students' performance on the state assessments. Several principals worried that although TEPP was potentially effective at filling in gaps in students' prior knowledge, the trade-off in time spent on grade-level skills disadvantaged students on the assessments. Likewise, although teachers in our study appreciated the personalized nature of TEPP, many still expressed concerns that their students were not "receiving the preparation they need to meet grade-level responsibilities" and would be "shocked" when they faced grade-level skills on subsequent standardized assessments. From the perspective of one of these teachers,

> We were all on board with this plan [to incorporate more grade-level skills] because we feel like exposure to grade-level content—even if you don't master it—really matters. But, when it comes to the [state test] and all these other things, it causes so much anxiety for them because they've never seen it and it deflates their confidence. They say things like, "I've never seen this, Miss!"

Teacher evaluation. These tensions were particularly consequential for teachers due to the existing high-stakes teacher evaluation system. Teacher concerns regarding student performance on the state grade-level assessments were intertwined with concerns regarding their own evaluation. Since TEPP did not focus exclusively on grade-level content, teachers questioned the fairness of being evaluated based on these tests. They argued that the test scores were neither an accurate reflection of instructional quality nor the quality of TEPP. Additionally, in traditional accountability systems, teachers are typically held accountable for achievement (or achievement growth) among students for whom they are clearly responsible. However, within the TEPP model, teachers work with all students across multiple grades, disrupting traditional notions of teacher evaluation. The district's response was to hold all teachers accountable for learning among all students in the program. Although assigning students to multiple teachers within the same year and subject is not a core element of most personalized models, reform groups are increasingly urging schools to consider eliminating the "teacher of record" designation so that on any given day, students can access the teachers most able to meet their individual learning needs (see Rodel Foundation, 2014).

Teachers reported mixed feelings about sharing responsibility for student learning within the high-stakes, value-added evaluation system. Many teachers viewed the communal approach in a

positive light, particularly those who were members of strong teacher teams. As one such teacher stated, "We love working as a team and are confident in each other's abilities, so in that sense, we are not concerned [about the evaluation results]." Similarly, another teacher shared that she preferred to "sink or swim" together, while a third felt that her team was so strong that improvements in student achievement were within reach. Others claimed that they trusted their teams, with one teacher stating that if "everyone does their job, we will be fine; in this building we work together and support each other." In general, schools with teaching teams that had already established strong interpersonal relationships and trust prior to TEPP seemed better able to cope with the new pressures of the competing logics. It appeared that collaboration and teamwork were already built into these teachers' schematic constructs about the role of teaching, which allowed them to better make sense of the model's new demands compared to newer teachers and those who had yet to establish strong interpersonal relationships with one another.

In contrast, other teachers interpreted the new approach to teacher accountability in a negative light. Key for this group of teachers was a sense of lost autonomy, with several lamenting that they "no longer had their own students." These teachers made sense of their roles in a more traditional, individualistic manner, where rewards and sanctions were based on their direct effects on student achievement. TEPP disrupted their sensemaking by diluting their direct control over whom and what they taught. For example, as one teacher reported, "Teachers are used to being captains of their own ship, but when you're doing [TEPP], you're part of a team, and you have to be able to leave your ego at the door." Another teacher felt "uncomfortable owning the results of all 200 students," particularly because they did not always see the same students every day. Several stated quite bluntly that they simply had little faith in their colleagues. Of the teachers who shared this sentiment, the dual, conflicting demands of the teacher accountability system and TEPP forced them to evaluate each other in a new light. Some looked at each other with more scrutiny, judging factors such as credentials or teaching experience as potential, preemptive explanations for their inability to meet shared accountability standards. For instance, one teacher questioned the fairness of evaluating teacher quality collectively when not all teachers on the team were certified in mathematics. The same teacher, however, felt self-conscious about their own contributions to the teaching team, claiming they felt "guilty" if they did not manage to "get students to understand the material."

Compromise: Navigating the Tensions

Organizational reconciliation. Over the course of the 3-year implementation, TEPP program staff and school administrators became more attuned to these competing logics and sought to reconcile them through a variety of organizational solutions. During the first 2 years of implementation, TEPP programmers offered to incorporate "floors" and "ceilings" into the assignment algorithm to ensure students did not receive content too far above or below their assigned grade level. TEPP also offered districts a more focused "test-prep" period during the weeks prior to high-stakes state assessments to prioritize grade-level material. In the district we studied, the incorporation of floors and ceilings varied among schools, across semesters, and even across grade level cohorts during the first 2 years of implementation. However, a floor of 3 years below grade level was imposed at all five schools. In addition, during the 1st year of implementation, two out of the five participating schools requested a test-prep period of 5 to 6 weeks in order for students to work exclusively on grade-level content. In the 2nd year, four of the five schools opted for this test-prep period, indicating increased concerns about student performance tied to the grade-level assessments.

In the 3rd year of implementation, TEPP responded to these concerns by providing more structure to the disparate floors and ceilings and creating a set of strategies to pilot across their portfolio of schools. In addition to the traditional test-prep window approach, the program developed two other strategies. In one new approach, grade-level floors were gradually introduced over the school year to avoid an abrupt shift to grade-level skills at the end of the year. Three district schools employed this model during the third year of TEPP implementation. In another new approach, students began each unit with grade-level skills and were only given practice with below-grade content if they were missing the foundational skills necessary to grasp specific grade-level content. Two district schools employed this model.

In addition to incorporating grade-level skills into the TEPP assignment algorithm, by the 3rd year of our study, school administrators were experimenting with site-specific organizational changes to support the incorporation of grade-level content. All schools offered some form of supplemental grade-level math instruction for students, including Saturday or after-school tutoring sessions, intervention periods, or even special test-prep days during which students devoted the entire 90-minute block to grade-level content instead of TEPP. Some schools used program data to tailor these extra interventions to students' needs, but these activities generally took place outside of and in addition to TEPP and were undertaken with the specific goal of helping students perform better on the state assessments. The addition of these structural supports aligns with suggestions in the literature that building separate, peripheral structures is necessary to address different environmental demands under the same roof (Honig & Hatch, 2004).

Individual reconciliation. Teachers made sense of these organizational attempts at reconciliation in a number of ways. Many teachers felt relieved by the incorporation of at least some grade-level content into the TEPP model, with one teacher claiming that it helped to determine which skills to focus on and clarified how far behind their students were academically. Earlier introduction of grade-level content during the third year of implementation also provided a level of reassurance to students, with some teachers claiming that it reduced the stress previously experienced by students in Years 1 and 2 when they were forced to abruptly leap across multiple grade levels in the weeks leading up to the state assessments.

Despite the increased comfort that came with knowing the earlier switch to grade-level material would prepare students for the state assessment, some teachers claimed that the negative consequences of turning "on" and "off" personalized and grade-level

content still remained under the new iterations of the model. One negative consequence was that high-achieving students were pulled back to content they had already mastered, finishing all their tasks early during grade-level review sessions. On the other end of the achievement distribution, students who were working on content below grade level would often become restless, distracted, and disheartened when moved up to grade-level material. Teachers argued that bouncing students back and forth from personalized to grade-level content not only confused them, but affected their self-esteem, increased their anxiety levels, and led them to become more dependent on teachers for instruction and assistance. One teacher remarked that she became aware when the platform started giving students grade-level content "because they stopped turning in their homework; they don't understand it."

An important takeaway from this reconciliation process is how these competing logics and subsequent compromises affected teachers' loci of control. How could they differentiate instruction to "meet students where they are" while simultaneously preparing students for assessments far beyond their current skill level? Staff members and students rarely held the TEPP platform responsible for this conundrum: They more often hoped that the accountability system would change to accommodate personalization, rather than see personalized instruction yield to the accountability logic. A few teachers even suggested prioritizing the implementation of the personalized learning program with fidelity over improving state test scores. One such teacher described the mixed signals she was receiving:

> I wish we were using it the way that it's supposed to be used. I've been saying that the whole time. Because, you know, all we're taught is differentiation and differentiation and then we have a program that actually, you know, develops a learning plan for each individual kid. And it's like, "Oh, no no. They actually have to be doing the same thing." But it's like, "Wait, that doesn't make any sense."

This teacher was not alone in feeling like the modified TEPP format was neither supporting growth the way it was intended nor supporting the demands of the accountability environment. As summarized by a different teacher frustrated with being pulled in different directions, "Screw the test! If you want the kids to grow, then let it work. You have to decide what you want—do you want growth or test scores?" Despite concerns about the compromised TEPP format, teachers were quick to remind us that the failure to include any grade-level content would have negative ramifications for both students and teachers.

Importantly, the delicate balance between personalized and grade-level content colored teachers' opinions about the sustainability of TEPP in the district. Overall, teachers who wanted to implement TEPP as designed, rather than its compromised format, sought more time to demonstrate that it worked. These teachers described how, under the accountability environment, there was a finite window of opportunity to prove the effectiveness of any new program or intervention. As described by one teacher, "The school district switches math programs every 3 years and never lets it run the whole cycle. I've been here 21 years, and I've been through probably six math programs." Indeed, we learned that in the year following our study, two of the five participating schools had discontinued the program and all schools had dropped TEPP by the end of the 2018–2019 school year.

Discussion

Scholars have long argued that policy incoherence is linked to ineffective school reforms (Cohen, 1982; Fuhrman, 1999; Shulman, 1983). Thus, whenever disparate policies with conflicting institutional logics are introduced into one organizational setting, educators, leaders, and other programmatic staff must continually negotiate what it means to abide by both logics. Our case study demonstrated how this process played out in an urban, low-performing school district. We observed how stakeholders' initial excitement for personalization through TEPP transformed into the chaotic adoption of additional responsibilities and creative ways of manipulating TEPP materials to ensure that students were still exposed to grade-level content prior to the state assessments. Although teachers and leaders expressed clear support for TEPP, they ultimately still expected their students to meet grade-level expectations, thereby defaulting to accountability logic to judge success.

Honig and Hatch (2004) argue that "crafting coherence" in policy implementation involves the identification of school-wide goals and strategies, as well as subsequent decisions about whether to buffer or bridge the school from or to external demands. With the five schools discussed here, buffering from institutional accountability pressures was nearly impossible. Rather, we observed that all five TEPP schools added peripheral structures to simultaneously meet the goals and strategies of both technology-enabled personalization and accountability. This strategic decision to partially engage in grade-level standardized test preparation and selectively "turn on" the tech-enabled personalization programming was an attempt for schools to signal their substantive commitment to both. Despite the addition of these peripheral structures, the competing demands of each logic undermined student academic improvement by both definitions. In our study, we witnessed how this manifested in a disruption of teacher sensemaking and an incoherent implementation process.

Limitations

Our findings should be interpreted in light of three specific limitations. First, our data come from a single high-poverty school district that had traditionally struggled on state-mandated grade-level assessments. It is possible that more-affluent, higher-performing districts and schools would not experience the same accountability pressures, and would perhaps feel that they had the freedom to experiment with personalized models without accountability sanctions. Interestingly, however, higher-performing districts in the U.S. have generally been *less* likely to adopt technology-enabled, personalized learning platforms. The reasons for this are beyond the scope of our study, but the question of which schools adopt similar models (and why) clearly deserves further scrutiny.

A second consideration is that this study examined the implementation of only one technology-enabled personalization model. It is unclear how our findings would vary across platforms with different components and expectations. Additionally, this particular technology-enabled platform focused exclusively on mathematics in the middle grades. Our findings may well have differed with an intervention that focused on other subjects or earlier grades that did not participate in the accountability system. Finally, although student focus groups were part of our study, we did not directly ask students how—if at all—they experienced the tensions between personalization and accountability.

Conclusion

Overall, we found that the delicate navigation of priorities, tasks, and technologies left teachers to their own devices to reconcile these tensions in their classrooms. We observed both disheartening and hopeful signs in this regard. Many teachers at these urban, high-poverty schools simply did not have the time to continually engage in the process of sensemaking, nor the resources to create and uphold peripheral structures to reconcile discrepant institutional demands. This lack of capacity may contribute to the inability of teachers and school leaders to execute coherent implementation plans for policies that clash with existing institutional logics, which may result in decisions to prematurely eliminate programs and innovations. More positively, some teachers and leaders shared creative strategies for reconciliation that may provide insight for generating and refining future models. These practitioner efforts could be interpreted as early signs of coherent adaptations. With the growing investment in technology-enabled personalization, researchers and product designers would do well to further explore the ways in which individuals are adapting to this new environment to inform future efforts. Additionally, future researchers should pay particular attention to how the navigation of these institutional tensions may be influencing the sustainability of school improvement efforts.

NOTE

Iris Daruwala is now at Social Policy Research Associates in Oakland, CA.

Support for this study was provided by the U.S. Department of Education, Institute of Education Sciences. All opinions and conclusions are solely those of the authors, and do not necessarily reflect the views of the funding agency.

REFERENCES

Bill & Melinda Gates Foundation. (2014). *Early progress: Interim research on personalized learning*. Author. Retrieved from http://k12education.gatesfoundation.org/student-success/personalized-learning/early-progress-interim-report-on-personalized-learning-rand/

Burnette, D. (2017, November 7). States take steps to fuel personalized learning. *Education Week*.

Cohen, D. K. (1982). Policy and organization: The impact of state and federal educational policy on school governance. *Harvard Educational Review*, *52*, 474–499.

DiMaggio, P. J., & Powell, W. W. (1991). Introduction. The new institutionalism in organizational analysis. In P. J. DiMaggio & W. W. Powell (Eds.), *The new institutionalism in organizational analysis*. University of Chicago Press.

Friedland, R., & Alford, R. R. (1991). Bringing society back in: Symbols, practices, and institutional contradictions. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis*. University of Chicago Press.

Fuhrman, S. (1999). The new accountability. *CPRE Policy Briefs.* Retrieved from https://repository.upenn.edu/cpre_policybriefs/14

Glazer, J. L., Massell, D., & Malone, M. (2018). Charter schools in turnaround: Competing institutional logics in the Tennessee Achievement School District. *Educational Evaluation and Policy Analysis*, *41*(1), 5–33.

Gross, B., & DeArmond, M. (2018). *Personalized learning at a crossroads*. Center on Reinventing Public Education. Retrieved from https://www.crpe.org/sites/default/files/crpe-personalized-learning-at-crossroads.pdf

Hallett, T. (2010). The myth incarnate: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review*, *75*(1), 52–74.

Haveman, H., & Gualtieri, G. (2017). Institutional logics. In R. Aldag (Ed.), *Oxford research encyclopedia of business and management*. Oxford University Press.

Herold, B. (2016, March 9). Facebook's Zuckerberg to bet big on personalized learning. *Education Week*.

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, *33*(8), 16–30.

Hyslop, A., & Mead, S. (2015). *A path to the future: Creating accountability for personalized learning*. Bellwether Education Partners.

Lizardo, O., Mowry, R., Sepulvado, B., Stoltz, D. S., Taylor, M. A., Van Ness, J., & Wood, M. (2016). What are dual process models? Implications for cultural analysis in sociology. *Sociological Theory*, *34*(4), 287–310.

Louis, M. R. (1980). Surprise and sense making: What newcomers experience in entering unfamiliar organizational settings. *Administrative Science Quarterly*, *25*, 226–251.

Lowenhaupt, R., Spillane, J. P., & Hallett, T. (2016). Education policy in leadership practice: "Accountability talk" in schools. *Journal of School Leadership*, *26*(5), 783–810.

Martinez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgements, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment*, *14*(78), 78–102.

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (Vol. 41). SAGE Publications.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340–363.

Meyer, J. W., & Rowan, B (2006). Institutional analysis and the study of education. In J. W. Meyer & B. Rowan (Eds.), *The new institutionalism in education* (pp. 1–13). State University Press of New York.

Molnar, M., & Herold, B. (2018, November 6). Are companies overselling personalized learning? *Education Week*.

Rodel Foundation. (2014). Personalized learning school models & educator roles. Author.

Rosch, E. (1978). Principles of categorization. In B. B. Lloyd & E. Rosch (Eds.), *Cognition and categorization*. Erlbaum Associates.

Rowan, B. (2006). The new institutionalism and the study of educational organizations: Changing ideas for changing times. In H. D. Meyer & B. Rowan (Eds.), *The new institutionalism in education* (pp. 15–32). State University of New York Press.

Scott, S. M. (1995). *Institutions and organizations*. SAGE Publications.

Shulman, L. S. (1983). Autonomy and obligation: The remote control of teaching. In L. S. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy* (pp. 484–504). Longman.
Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387–431.
Strauss, C., & Quinn, N. (1997). *A cognitive theory of cultural meaning* (Vol. 9). Cambridge University Press.
Vaughan, D. (1996). *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. University of Chicago Press.
Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*, 1–19.
Weick, K. E. (1995). *Sensemaking in organizations* (Vol. 3). SAGE Publications.
Weick, K. E. (1996). Fighting fires in educational administration. *Educational Administration Quarterly*, *32*(4), 565–578.
Wolf, M. A. (2010). *Innovate to educate: System [re]design for personalized learning*. Software & Information Industry Association.

AUTHORS

IRIS DARUWALA, PhD, is a Senior Associate at Social Policy Research Associates, 1333 Broadway Suite 310, Oakland, CA 94612; *ih2291@tc.columbia.edu*. Her research examines the organizational sustainability of education reforms, the role of philanthropy in shaping education policy, and the relationship between education policies and social stratification.

SHANI BRETAS, MA, is a PhD student in the Department of Education Policy and Social Analysis and a Senior Researcher at the Consortium for Policy Research in Education at Teachers College, Columbia University; 525 W. 120th St. New York, NY 10027; *ss4290@tc.columbia.edu*. Her research focuses on evaluating programs and policies designed to increase access to high-quality K-12 education and examining the structural and organizational conditions that shape program implementation.

DOUGLAS D. READY, PhD, is an Associate Professor of Education and Public Policy and the Director of the Consortium for Policy Research in Education at Teachers College, Columbia University, 525 W. 120th St. New York, NY 10027; *ddr2111@tc.columbia.edu*. His research explores the links between education policy, social policy, and educational equity, with a particular focus on how contemporary policies moderate or exacerbate socio-demographic disparities in student outcomes.

Manuscript received November 22, 2019
Revision received April 18, 2020
Accepted August 9, 2020