

# Reimagining Teacher Evaluation and Development as Mutual Networks for Agency and Growth

Baugh, A., & Ravdin, D.

American University

In the ongoing national conversation about teacher quality and retention, education leaders and policymakers have invested considerable effort in expanding pathways into the teaching profession. These efforts—ranging from alternative certification routes to financial incentives—aim to diversify the teacher workforce and alleviate widespread staffing shortages, especially in high-need urban districts. However, such recruitment strategies have rarely been matched with parallel reforms in the systems that shape teachers' everyday professional lives: teacher evaluation and professional development (PD). As a result, many of the educators recruited through these pipelines, particularly Black and Brown teachers, enter school environments that fail to sustain their growth, affirm their identities, or support their long-term professional trajectories. We argue for a fundamental reimagining of teacher evaluation and PD as interconnected developmental networks—systems that center reflection, teacher identity, mentorship, and culturally responsive practice. Drawing from transformative learning theory, constructive development theory, and insights from two distinct urban school contexts: a middle school in Washington, D.C., and a charter middle school in the Bronx, we utilize educator perspectives on formal and informal evaluation processes, how they interpret the feedback they receive, and whether those experiences translate into meaningful professional development to develop a conceptual model that positions teachers not as passive recipients of feedback and training, but as co-constructors of their growth. Our framework emphasizes the importance of agency, community, and culturally responsive feedback cycles, challenging the dominant paradigm that treats evaluation and PD as discrete, top-down mechanisms for monitoring compliance and offers a model for how districts and schools can begin to design evaluation and PD systems that truly support equity, agency, identity, and long-term educator wellbeing.

*Keywords:* Teacher development, teacher evaluation, coaching, teacher collaboration, collective efficacy

## Introduction

In many schools, teacher evaluation and professional development remain siloed, top-down processes characterized by standardization, surveillance, and deficit assumptions. Teacher evaluation systems—often anchored in technocratic rubrics and observation checklists—fail to capture the complexity of teaching, reduce dynamic classroom practice to segmented snapshots, and too often serve as compliance tools rather than meaningful sources of growth (The New Teacher Project, 2009; Kraft & Gilmour, 2017). Likewise, professional development is frequently generic, one-size-fits-all, and disconnected from the lived realities of teachers' classrooms. For educators working in Title I schools or serving historically marginalized students, PD often focuses more on classroom management and technical fixes than on culturally responsive pedagogy or reflective inquiry (Campbell, 2023; Darling-Hammond et al.,

2017).

The impact of these structures is particularly acute for teachers of color. Research consistently shows that Black educators receive lower observation scores than their white counterparts, even when controlling for student achievement and school context (Grissom et al., 2021; White, 2020). These inequities in evaluation not only shape teachers' access to development opportunities and advancement but also contribute to burnout, demoralization, and high attrition rates—ultimately undermining the very goals of diversity and retention initiatives (Bireda & Chait, 2020; Ingersoll et al., 2019). At the same time, teachers of color report taking on disproportionate responsibility for school culture, student support, and extracurricular programming—forms of labor that are rarely reflected in evaluation systems or rewarded in PD structures (Bristol, 2018; Ravitch, 2014).

Compounding the issue is the fact that both

evaluation and PD are typically framed through a white, middle-class lens of what “effective teaching” looks like, often overlooking culturally sustaining practices and discouraging authentic teacher voice. Teachers are expected to conform to norms codified by observation tools that prioritize compliance, performative engagement, and visible behaviors over relational, inquiry-driven, or student-empowering practices. The result is a system that stifles innovation, erodes intrinsic motivation, and alienates many of the very educators’ schools aim to recruit (Campbell, 2023; Paris & Alim, 2017).

## Literature Review

To reimagine teacher evaluation and professional development as interconnected developmental networks, we ground our analysis in two complementary frameworks: Transformative Learning Theory (Mezirow, 1991) and Constructive-Developmental Theory (Kegan, 1994). These theories of adult development help explain why current evaluation and PD systems often fail to support meaningful teacher growth and what a more teacher-centered, equity-focused alternative might require. Transformative Learning Theory posits that adult learning is not merely about acquiring new strategies, but also involves a fundamental reevaluation of one’s beliefs, values, and assumptions. As Mezirow (1991) argues, adults learn best when they critically reflect on their lived experiences, engage in discourse, and revise their “frames of reference” in ways that allow for deeper, more authentic engagement with their work. In the context of teaching, this means that growth does not emerge from top-down mandates or checklist feedback, but from experiences that prompt teachers to reflect on their practice, consider new perspectives, and construct new understandings. Yet most teacher evaluation systems are designed to reinforce compliance rather than encourage this kind of reflective transformation. Observation rubrics and summative ratings rarely provide the time, space, or support necessary for teachers to critically examine their pedagogical approaches or to explore the intersection between their professional practice and personal identity, particularly in schools serving racially and linguistically diverse communities (Brookfield, 1995; Cranton, 2006; Burciaga & Kohli, 2018). Simultaneously, Constructive-Developmental Theory has decades of empirical evidence supporting its suggestion that adults’ development occurs in a set progression that builds on development in previous stages to shape people’s meaning-making (referred to as “ways of knowing,” or WOK), resulting in rich cognitive and analytical capacities that shape adults’ orientations to learning, leadership, and contextualization. “Growth, according to constructive-developmental theory, is connected to increases in an individual’s cognitive, affective,

interpersonal, and intrapersonal capacities,” (Drago-Severson & Blum-DeStefano, 2019) or, as Kegan (1994) framed these WOKs: instrumental, socializing, self-authoring, and self-transforming. Leaders concerned with the development of adult professionals, such as teachers, can utilize these cumulative mentalities to build structures to enhance personal and interpersonal learning in their settings. As an adult learner transitions from a concrete understanding of themselves and their orientations (instrumental) to adopting the perspectives of those around them on themselves and their environment (socializing), the previous WOK becomes deeper and more nuanced. Meaning-making through socializing provides a more nuanced and comprehensive sense of the instrumental self and position. As these learners gain the capacity to incorporate those perspectives and reflect on others’ opinions to shape their own values and actions (self-authoring), they begin “engaging in conflict without feeling torn apart, thinking systemically, and enacting a clear vision” (Drago-Severson & Blum-DeStefano, 2019). This stage comes with drawbacks, however, as these closely held self-authored ideas may resist interrogation or expansion. “Because of this, self-authoring adults can appear less open to new ideas, especially to ideologies and perspectives that diametrically oppose their own” (Drago-Severson & Blum-DeStefano, 2019). For this reason, school leaders should prime their developmental environment to help learners reach the fourth stage, self-transforming, wherein these self-authored philosophies are coupled with a hunger for new perspectives and ideas to incorporate into one’s meaning-making. This stance promotes true cooperation and collaboration within a school community. Drago-Severson and Blum-DeStefano (2019) point out that Robert Kegan estimates only 8-11% of the US population reaches this stage. As such, the moment calls for more educators to have these capacities developed in their settings through mutual learning partnerships, structured reflection, and development that centers the aspirations and values of individual teachers.

## Methodology

This conceptual paper draws on insights from two urban middle schools, a traditional public school in Washington, D.C., and an arts-focused charter middle school in the Bronx, exploring how teacher evaluation and professional development systems are experienced in real-world contexts. Through environmental scans, semi-structured interviews, and reflective narratives, patterns of disconnection, bias, and untapped potential surface in these systems. These insights both confirm and deepen the theoretical critiques outlined in the literature review and help illustrate what a more equitable, teacher-centered model might look like.

At the D.C site, data were collected through one author’s work as an instructional coach within the

District of Columbia Public Schools (DCPS), where the multiple-measure IMPACT evaluation system governs instructional feedback. Semi-structured interviews were conducted with a classroom teacher, an instructional coach, and a principal to explore how feedback was delivered, interpreted, and used to support or hinder professional learning and growth for teachers. The Bronx site provided a complementary lens. There, the second author conducted interviews with four general education teachers, a special education teacher, and a fellow coach around the tools, practices, and feedback utilized for development and evaluation.

Though the systems differed in structure—DCPS’s centralized IMPACT model versus the Bronx school’s self-developed rubrics—the findings converged around three core themes: (1) the feedback was disconnected, cultural labor and identity work were unrecognized, and feedback was delivered inconsistently or felt biased.

## Findings

### Theme 1: Feedback is Disconnected from Professional Growth

Across both sites, educators described formal feedback structures as rigid, rushed, and misaligned with their actual development needs. In D.C., a veteran teacher noted, “I actually liked when I had the master educators better than having the administration in, because it [the process] can be biased... because the way I teach is more project-based...” The veteran teacher questioned how someone without teaching expertise in their subject could effectively evaluate them, stating, “How can you have someone who’s not on the teaching lens, evaluate me, not knowing the position?” They suggested that instructional coaches who are content experts should conduct evaluations instead. Although instructional coaches often offered more tailored feedback, the high-stakes nature of formal evaluations meant teachers prioritized “scoring well” in the domains of the rubric over deepening their practice.

Most of the teachers interviewed shared that any feedback they received after these monthly observations was focused on what a fourth-year teacher called “quick fixes and immediate teacher moves rather than long-term growth or development goals.” The teacher went on to provide a written example a day later: “after a brief 5 to 10-minute observation, administrators would sometimes leave a note on a Post-it or index card with suggestions like making a quick adjustment or paying closer attention to specific CFUs [Checks for Understanding]...Feedback was often centered on immediate changes rather than fostering a broader discussion about growth or development aligned with my teaching style.” Despite the specious appearance of soliciting input by having teachers complete self-evaluations, all of them agreed that the conversations were not collaborative or inclusive of personal goals. One teacher who said that her perspectives were asked for in discussions

complained that it included only growth areas that the administrators had unilaterally chosen as goals. “Sometimes there were administrators who considered or asked about my perspectives on the areas they identified. Occasionally, some would acknowledge improvements by saying, ‘See, that improved,’ while others would ask how certain strategies worked out. However, these instances were not frequent, and overall, my input was seldom solicited in a meaningful way.”

In the Bronx, a former teacher now on the leadership team discussed how there is a ceiling for the professional growth available for teachers based on the rubric, “the school doesn’t have a plan for how to support teachers once they hit a certain threshold. Not everybody can enter leadership once they’ve met the ‘meeting’ section of the rubric [in] all five categories, so it’s unrealistic to expect all of the teachers that are currently there to stay without a real development plan for how to support [them once] they’ve reached that point [in] their careers.” With this reflection, broader instructional feedback is more akin to a mechanical process that meets a rubric, rather than coaching teachers with a growth mindset as whole human beings and pushing their practice in areas such as critical consciousness, teaching emotionally dysregulated students, or improving content knowledge.

In classrooms, effective teaching practice requires effective feedback to build student mastery. Descriptive feedback, banal praise, and non-actionable quality judgments are not effective drivers of student development. Similarly, students will naturally feel less willing to advocate for themselves or take risks if they think that a teacher’s feedback is based on interpersonal feelings. Perceptions of student choice, facilitated autonomy, and two-way feedback mechanisms help to drive effective learning and teaching in the best classrooms. It is natural to assume that these same principles would apply when it comes to feedback to teachers from administrators, particularly those who are also employers. It is difficult to imagine a teacher feeling strong motivation for personal growth in a setting where they feel like that growth must take place around or directly in opposition to the practices that determine their future in the school.

### Theme 2: Cultural Labor and Identity Work are Invisible

When asked about what they perceived to be the largest impact that they have on their school, all of the educators cited elements that were not captured by the evaluation documents. The teacher who has since joined the LT was certain of hers:

I think my largest positive impact is definitely [that] my relationship with students once they leave my classroom gets stronger, but also they value what they learn with me because it applies to their real lives. I think that when I share real aspects of adulthood or myself

growing up, it became clearer that I had a larger impact because students could relate to me more and see me as more than just a teacher...I don't think the rubric captures all of that.

An ELA (English & Language Arts) teacher also felt her largest impact was not reflected in her evaluative documents. "I believe my largest impacts were creating a growth mindset culture within my students, an openness and willingness to take risks in reading and writing, and fostering a love of reading in students. Many students left my class liking reading more than they had previously and seeing themselves as writers." While this impact was not measured on rubrics or referenced in her feedback, it was lauded by her fellow teachers: "I had colleagues point out that students continued to have the self-esteem and growth mindset that they felt was developed in my class. They sent students to me to show me their grades in the next grade and how they felt the culture created in my class contributed to that." Another classroom teacher also said she received cards from colleagues appreciating the impact she had on their students. Others pointed out that students themselves also appreciated elements that were never part of their formal evaluations, a criticism that was echoed by teachers measured by IMPACT.

Although the significant contributions these teachers made were not reflected in formal individual evaluations, they have had a profound impact on the school's growth in test scores, attendance, and high school acceptances, as well as a notable decrease in disciplinary infractions and conduct violations. There have also been large increases in interest and offerings of extracurricular activities, which is a keen indicator of the strong culture these teachers have built. Even so, the teachers feel that the appreciation for the impact they had on these elements is thin. "I think the impact was slightly recognized, always appreciated, but slightly recognized... Most promotional materials [for the school] were created from things that I did in my classroom, but none of the things I did outside of my classroom were highlighted."

### **Theme 3: Teachers Felt Valued and Saw Development Through Collaboration with Peers**

Despite the limitations of formal evaluations and current PD structures, both sites revealed a promising theme for teacher growth through informal peer networks and collaborative practices. These spaces often occur outside the formal evaluation and development system and serve as critical developmental networks for teachers to feel empowered, reflective, and to co-construct knowledge. In the Bronx, a classroom teacher said that more consideration of her goals came from her colleagues and coach than from her evaluators and observers. "I was not asked by those who are supposed to be my direct supervisors about my personal goals or desired areas for growth

during these conversations...There was little to no discussion about my aspirations, and my input on areas where I wanted to develop was rarely, if ever, solicited. However, I was asked about this from others in the building who were willing to help me improve." One of her colleagues, who had partnered with her on some of these goals, agreed that most of the driven, authentic growth came from colleagues, not observers or coaches: "The primary support I received in developing the identified areas came mostly from other teachers. When I reached out to them, they were helpful in providing practical advice on specific strategies... In a few instances, my coach provided more structured support by modeling specific activities, showing me how they should be delivered in the classroom, which helped me implement those strategies effectively." This could be an effect of the impersonal and arbitrary sense of the tools that teachers expressed. From these two examples, it is clear that the formal systems for teachers remove teacher voice from the development and thus promote a sense of support, value, and professional agency.

Simultaneously, in Washington, D.C., the principal also noted how the IMPACT evaluation system inadvertently promotes peer collaboration. The administrator noted that teachers become "more vulnerable with one another" within departments. Teachers willingly share lesson plans for feedback before formal evaluation, and an informal knowledge-sharing network builds from teachers who perform well in evaluations. The administrator continues, "If someone's done particularly well, well... that person becomes pretty popular because they're like, 'What did you do? Could I see your lesson plan?'" The administrator's comments reflect the growing problem within the practice of teacher evaluation and development. The administrator in D.C. observes the system naturally forming an informal network, which may highlight a strength of the multiple measure system, as the teachers in the Bronx felt they had to look outside their evaluation system for more consistent and targeted development in their personal growth areas.

### **Conclusion**

Considering Langley et al.'s (2013) critical features of fundamental change, the opportunity is ripe for overturning some of these implicit assumptions baked into the synergistic evaluation/development systems based on ersatz metrics that have become targets. By rearranging the process and services provided by coaches and augmenting the desired product of the thought partnership, the system for developing teachers gets a much-needed redesign. It also provides an improvement to an existing dynamic without needing significant structural upheaval because the problems are those of emphasis and targets, not of purpose or intent. That said, it does fundamentally depend that dynamic by centering the teacher and

reassigning the evaluation documents to metrics, rather than plans. Thus, the learning network can radiate out from them with expert guidance, rather than trickling down learning from arbitrary, non-reflexive documents. One can expect the result of upsetting these assumptions to be seen not only in the quality of teaching but also in nonacademic impacts of teachers on students and schools, teacher agency and assertiveness, and collaborative efforts due to building agency and psychological safety in a coaching relationship that centers teachers. Finally, one can expect that by moving from a segmented set of static targets to check off to a system of growth, exploration, and analysis, teachers will gain the skills to drive their development and impact their peers, sending reverberations of these impacts into the future. Providing the supports and metrics to teachers who have chosen targets will negate any possibility of Goodhart's Law skewing those impacts, can reframe our notion of teacher improvement from target practice to exponential growth, and can place the problem focus where it needs to be: on giving willing teachers the tools, perspectives, expertise, and support they need to grow in ways to benefit the entire school ecosystem and build collective efficacy to drive culture change. To paraphrase Ross Greene, teachers do well if they can, and do better if they center aspirations in goals, see their peers as resources, and drive their learning toward their aspirations.

## Discussion

School and district leaders can prime their environments for this development by building such developmental networks around the following recommendations: evaluation as a part of continuous learning, formalized peer mentoring, and center the teacher. These are elaborated upon in the following sections.

### Evaluation as a Part of Continuous Learning

Evaluation should be reframed as a measure of ongoing professional growth, not simply an annual verdict on teacher effectiveness. Rather than rely on isolated observations or narrow scoring rubrics, school leaders can embed evaluation within a cycle of reflective goal-setting, iterative feedback, and developmental coaching. To improve the formal observation's place in the learning cycle, instructional coaches, thought partners, and mentor teachers must all be part of the evaluative process, not just administrators, to ensure that feedback is personalized, actionable, and grounded in a shared understanding of pedagogy and content. This shift takes the evaluation from a snapshot held against a static, standardized metric to one that better honors teacher development and recognizes that growth is nonlinear, contextual, and autonomous. It also requires the assessors and observers to move from a fragmented approach that attempts to isolate the individual elements of teaching and speciously

combine them to a holistic perspective. This new stance recognizes the dynamic interactions and reinforcing elements within a teacher's focus area, enabling feedback that is actionable, reflexive, and growth-oriented.

### Formalize Peer Mentoring

Teachers consistently describe peer collaboration as the most valuable and affirming form of professional learning. School leaders can institutionalize these practices by building structured opportunities for teachers to observe one another, co-lead PD sessions, and engage in critical feedback cycles. Instructional rounds (City et al., 2009) and inquiry groups offer teachers the opportunities to engage in complex instructional challenges across classrooms and content areas. Networks offering disparate perspectives and the opportunity to share and digest observations push learners into self-transformative WOKs and provide inroads to highlight and spread innovative practices and methods, anchored by research and ground truth. These spaces must be protected by school schedules and supported with time, resources, and recognition. Peer coaching models that are grounded in psychological safety and affirmed identity offer rich opportunities for true differentiated development that general, hierarchical school-based PD fails to provide. Recognizing peer mentorship as both a form of instructional leadership and professional development strategies creates distributed pathways for mutual growth, especially in school districts whose systems marginalize that informal expertise (Kohli et al., 2021).

### Center the Teacher

Too often, teachers' goal-setting begins with what they want to improve or add to their practice. This ignores the fact that teachers' professional plates are already overflowing with explicit and implicit expectations, so further drawing from a finite capacity will result in giving short shrift to the type of introspective and collaborative work described here. Instead, developmental relationships should counterintuitively begin with deprioritizing aspects of a teacher's professional duties to make mental and temporal space for the analysis, innovation, and reflection to take place. Coaches, mentors, and other thought partners can begin by bucketing elements of a teacher's role into "checklist" items (those that require completion but where increased effort will detract from the high-value and high-leverage elements) and "effort" items. This values-based prioritization work still utilizes the mentor/coach's expertise—as neophyte teachers may not necessarily see hidden value in some elements—but also communicates a recognition that space needs to be made for improvement and implementation. This also facilitates next-level conversations wherein the experience in a partnership can highlight the network of interactive elements that orbit and undergird that teacher's individual goals, moving

them into a self-authoring stance to prepare for the self-transformative work that drives personal growth, ripples through networks of peers, and nurtures a developmental ethos for entire school ecosystems.

## References

- Bireda, S., & Chait, R. (2020). Increasing teacher diversity: Strategies to improve the teacher workforce. Center for American Progress. <https://www.americanprogress.org>
- Bristol, T. J. (2018). To be alone or in a group: An exploratory study of Black male teachers' experiences in two U.S. school districts. *Teachers College Record*, 120(6), 1-28.
- Brookfield, S. D. (1995). *Becoming a critically reflective teacher*. Jossey-Bass.
- Burciaga, R. & Kohli, R. (2018). Disrupting whitestream measures of quality teaching: The community cultural wealth of teachers of color. *Multicultural Perspectives*, 20(1), 5-12. DOI: 10.1080/15210960.2017.1400915
- Campbell, S. L. (2023). Professional development or professional surveillance? A critical examination of how professional learning is experienced by Black teachers. *Urban Education*, 58(3), 451-478. <https://doi.org/10.1177/00420859221134884>
- City, E. A., Elmore, R. F., Fiarman, S. E., Teitel, L., & Lachman, A. (2018). *Instructional rounds in education: A network approach to improving teaching and learning*. Harvard Education Press.
- Cranton, P. (2006). *Understanding and promoting transformative learning: A guide for educators of adults* (2nd ed.). Jossey-Bass.
- Darling-Hammond, L., Hyler, M., & Gardner, M. (2017). Effective teacher professional development. <https://doi.org/10.54300/122.311>
- Drago-Severson, E. & Blum-DeStefano, J. (2019). A developmental lens on social justice leadership: Exploring the connection between meaning making and practice. *Journal of Educational Leadership and Policy Studies*, 3(1).
- Greene, R. (2008). Kids do well if they can. *Phi Delta Kappan*, 90(3), 160-167.
- Grissom, J. A., Blissett, R. S. L., & Mitani, H. (2022). Evaluating principals: Race and gender bias in principal performance ratings. *American Journal of Education*, 128(2), 165-200. <https://doi.org/10.1086/717143>
- Ingersoll, R., May, H., & Collins, G. (2019). Recruitment, employment, retention and the minority teacher shortage. *Education Policy Analysis Archives*, 27, 37. <https://doi.org/10.14507/epaa.27.3714>
- Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Harvard University Press.
- Kohli, R., Pizarro, M., Garcia, L.-G., Kelly, L., Espinoza, M., & Cordova, J. (2021). Critical professional development and the racial justice leadership possibilities of teachers of colour in K-12 schools. *Professional Development in Education*, 47(1): 89-101. <https://doi.org/10.1080/19415257.2020.1814387>
- Kraft, M. A., & Gilmour, A. F. (2017). Revisiting the widget effect: Teacher evaluation reforms and the distribution of teacher effectiveness. *Educational Researcher*, 46(5), 234-249. <https://doi.org/10.3102/0013189x17718797>
- Langley, A., Smallman, C., Tsoukas, H., & Van de Ven, A. H. (2013). Process studies of change in organization and management: Unveiling temporality, activity, and flow. *Academy of Management Journal*, 56(1), 1-13. <https://doi.org/10.5465/amj.2013.4001>
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.
- Paris, D., & Alim, H. S. (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.
- Ravitch, D. (2014). Hoaxes in educational policy. *The Teacher Educator*, 49(3), 153-165. <https://doi.org/10.1080/08878730.2014.916959>
- The New Teacher Project (Tntp). (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. <https://tntp.org/publications/view/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness>
- White, T. (2020). The hidden curriculum of teacher evaluation: How race shapes perceptions of feedback and instruction. *Journal of Teacher Education*, 71(2), 163-175. <https://doi.org/10.1177/0022487119841882>