Assessment and Accountability Policies:

A Critical Investigation of Rhetoric and Experience

_____

A Dissertation

Presented to

The Faculty of the School of Education and Human Development

University of Virginia

_____

In Partial Fulfillment

of the Requirements for the Degree Doctor of Philosophy

By

Sarah Beach

May 2023

ABSTRACT

Since the 1980s, public schools in the United States have been dominated by a system built around standards, standardized assessment, and high-stakes accountability. In more recent years, there have been movements to encourage public schools to focus more on 21st century skills. Currently in the Virginia system, there is a pressure to operate under both modalities with an accountability system based on the Standards of Learning assessed through standardized testing while also preparing graduates for their post-secondary life through the 5 C's: collaboration, critical thinking, creative thinking, communication, and citizenship skills. This three-paper dissertation critically considers the system in Virginia through: 1) a critical policy analysis of Virginia's assessment system and the conflicting assessment policy messages produced by the political process, (2) the resolution of the resulting policy conflict by differences in interpretation by policymakers at the state level and teachers in one school division, and (3) a consideration of how teachers measure success today, envision success reimagined, and are constrained by the current assessment and accountability system. The work concludes by considering the potential implementation of new systems of assessment and accountability that are more de-centralized, where school can be a humanizing place for students and teachers, and students can more authentically learn about themselves and the world around them.

Keywords: standards, assessment, accountability, performance-based assessment

Department of Leadership, Foundations, and Policy
School of Education and Human Development
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, "Assessment and Accountability Policies: A Critical Investigation of Rhetoric

and Experience," has been approved by the Graduate Faculty of the School of Education and

Human Development in partial fulfillment of the requirements for the degree of Doctor of

Philosophy.

_____

Walter Heinecke, Ph.D., Chair

_____

Katrina Debnam, Ph.D.

_____

Diane Hoffman Ph.D.

_____

Stanley Trent, Ph.D.

_____

Date

DEDICATION

To my students, past and present, for all you have taught me. Teaching you has been one of life's

greatest joys. This is for you, with the intent to create systems that support all students as they

pursue their dreams and make them come true.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

List of Tables

List of Figures

Linking Document: Introduction

This introduction serves as a linking document to provide the conceptual and theoretical linkages among all three manuscripts of this three-paper dissertation. Here, I tie the papers together by providing a brief historical overview of the standards, assessment, and accountability policy system in operation in the U.S. since the 1980s that will illustrate some of the controversies that have evolved to set the background for the following papers: (1) a critical policy analysis of Virginia's assessment system and the conflicting assessment policy messages produced by the political process, (2) the resolution of the resulting policy conflict by differences in interpretation by policymakers at the state level and teachers in one school division, and (3) a consideration of how teachers measure success today, envision success reimagined, and are constrained by the current assessment and accountability system. One of the overarching goals of the project is to address the ramifications of the current system for Virginia policymakers considering the potential of implementation of new systems of assessment and accountability.

Many of the policy debates within the field of education, when distilled to their core, can be traced back to divergent values (e.g., equity, excellence, choice, efficiency, etc.) and goals for education itself. Educational scholars and citizens of the United States have long debated the purpose of schooling, their goals often, but not exclusively, falling into three dedications: preparing equal citizens for democratic society, preparing people for the workforce, and/or teaching skills for economic mobility (Labaree, 1997). Notably missing from these goals is the more holistic concept of preparing students for life-long learning and their own human development. Of interest in the current debate around large-scale assessment is whether the content being assessed is aligned with the current multiple goals of the education system at large, and if so, *whose* goals are they?

The current era of assessment policy has been influenced by a political movement that was started in the 1980s known as the standards, assessment, and accountability movement. Through the rise of standardized testing following the publication of the National Commission on Excellence in Education's (NCEE) "A Nation at Risk" (NCEE, 1983), national and state policymakers have been committed to a standards-based curricular reform movement, tied to an accountability regime centered on standardized, high-stakes testing. This movement has shaped federal and state policies since that time in various forms.

Alternatively, students, teachers, and scholars of education have a long history of contesting this regime – asserting it narrows the curriculum, with social studies, science, art, and physical education being marginalized, while pushing teachers to teach to the test (Aurora Institute et al., 2021; Booher-Jennings, 2005; Henderson et al., 2022; Newmann et al., 1997; Smith et al., 1990). When the values reflected in governmental policies and the values reflected within the profession and community contradict each other, the conflict is often worked out in how the policies are implemented (Berman and McLaughlin, 1976; Odden, 1991; Viennet & Pont, 2017). Guba (1984) asserted that policy-in-intention often differs from policy-in-experience; however, policymakers rarely have opportunities to explore these differences.

The standards and accountability movements have been influenced by federal policy incentives, but they are enacted in a decentralized state system. Therefore, it is essential to understand the context of state standards and accountability reforms, as well as how those reforms have influenced the resultant emphasis on various forms of student testing. While many perceive machine scorable exams as the most valid, unbiased, and efficient form of assessment (Wilson, 2004), the history of standardized testing and the research results across the last forty

years have caused many scholars to question this objectivity (Au, 2016; Culpepper-Ofori, 2012; Gould, 1996; Randall, 2021, Randall et al., 2022).

Critical perspectives can help us to examine the facets of these arguments and determine for who and by whose standards education is successfully attaining outcomes. Writing about social science methodology, Flyvbjerg (2001) encourages all social science research to place power at the core of such analysis. He asserts the driving question of all social science research should be "what should we do?" (Flyvbjerg, 2001, pp.136-7).

Like all policies, assessment policy has been influenced by multiple, and often conflicting, values and by political power. In the recent era, the value of efficiency has been centered within the narrow framing of standards, accountability, and assessment policy, while competing policy values have included liberty/choice, equity, and quality/excellence (Heinecke, et al., 2003). Over the past fifty years or so, when researchers, parents, and policymakers consider the success of American public schools, it has become commonplace to prioritize only scores on reading and math assessments as the indicators of success. In an efficiency-oriented political culture of standards, assessment, and accountability, numerical data is accepted as scientifically objective. It's often quipped, '*numbers don't lie,*' but a principle known as Campbell's law warns against the risks of using a single number as a basis for telling a whole story: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it was intended to monitor" (Campbell, 1979, as cited in Nichols & Berliner, 2007, pp. 26-27). And while high-stakes accountability assessment systems often rely on the scores of a few end-of-year tests, people generally agree that schools should be responsible for teaching students much more than what is measured on those assessments.

While explained in more depth in paper one, a quick review of the definitions used throughout this dissertation is helpful here. The focus of this dissertation is the competing systems of *standardized* versus *performance-based assessments.* Standardized assessments may be in multiple choice or technology-enhanced formats, but generally are answered with one correct answer. Performance-based assessments focus more on process and require students to complete tasks or solve problems that can be addressed in any number of ways. Scoring for these assessments generally uses rubrics.

Accountability systems assign stakes to these assessments and can be at various levels from the individual student all the way up to the federal-level. When low stakes are assigned, the assessment may just be used as a data point. When high-stakes are assigned, students may need a particular score to graduate, and schools may need a particular pass rate in order to be granted accreditation. Aggregate scores are often used to rate schools, and punitive sanctions, such as state control of curriculum, often occurs when school pass rates are too low. Both standardized and performance-based assessments can be high- or low-stakes.

While critical reflection is generally a good practice, it is imperative when high-stakes decisions are made based on particular data points. Within such systems, it is important to consider who these systems are framed around and for whom they are intended. The issue of race and racism is prescient in the field of assessment policy. Love (2019) reminds us that while the narrative around accountability asserts that it is protecting our students and their futures, during this regime "we stopped protecting dark students' potential, if we ever had" (p.122). Kendi (2016) furthers this line of thinking with his declaration that "standardized tests have become the most effective racist weapon ever devised to objectively degrade Black minds and legally exclude their bodies" (paragraph 13). The history of assessment to be covered in paper one

supports this assertion. While the No Child Left Behind Act of 2001 (NCLB) made claims that the use of accountability was to help historically marginalized students, many scholars refute this, noting that the color-evasive, also referred to as colorblind, policies actually did, and perhaps even intended to do, the opposite (Au, 2016; Hollingworth, 2009; Yilmaz et al., 2022). Further, NCLB's continuing legacy perpetuates inaccuracies about achievement by students of color and continues to replicate a system in which those with access, privilege, and power are said to gain opportunity through merit, while those with less opportunities are punished and given far fewer chances to be successful (Au, 2016; Mandery, 2022; Yilmaz et al., 2022).

Love (2019) encourages a reframing of accountability outside of testing and even academic achievement, but instead "accountability that focuses on justice, love, healing, and restoring humanity. Educators, and especially those with privilege, must be responsible for making sure dark children and their families win" (Love, 2019, p.122). This leads us to the critical framing of this dissertation that includes an examination of equity, including racial equity, in standards, assessment, and accountability policy.

Policy Frameworks for Researching Assessment

While policy and research are often considered value-neutral, research on policy indicates this is simply not the case (Flyvbjerg, 2001; Lincoln & Guba, 1985; Rein, 1976). Policies themselves are created within a framework dominated by values, even if the values are not outright acknowledged. Assessment policy, like all policy, has always been shaped by divergent values existing within the political culture of the U.S.

Rein (1976) asserts that the societal goals and purposes of policies are constantly changing and that research supporting or contradicting those policies is always used, and often

mis-used, within the political realm. Therefore, we need a conceptual framework that allows us to critically analyze the assessment policy system.

Standardized testing in the current era is tied to efficiency, with machine-scorable tests allowing students, teachers, schools, districts, states, and countries to be compared easily. It is easy to see, however, that this efficiency runs counter to the goals of quality in assessment, getting a clear picture of what students know that takes more sophisticated forms of assessment but also more time and resources. Values of equity, quality, and excellence have in turn impacted policies, but the value of efficiency has largely dominated national assessment policies (Heinecke et al., 2003). The nation is in a moment of historical examination, and some states have been experimenting with alternatives to high-stakes standardized testing that promote excellence and equity over efficiency. In this dissertation, I will interrogate the varying and conflicting value positions behind high stakes accountability-based standardized assessment and emergent alternative assessment approaches, especially within the Virginia context.

The goal of this dissertation research is to analyze written and implemented policies to determine how the current assessment and accountability system in Virginia is working. It will also compare standards, assessment, and accountability policy-in-intention from policymakers' perspectives and policy-in-experience from the perspective of teachers interpreting and then implementing those policies. Several critical perspectives will be used to make these determinations.

Researcher as Instrument Statement

As a part of this critical work, I acknowledge the importance of identifying my positionality to the research. I have personal teaching experience in the state of Virginia, and while I have my own opinions and perspectives about assessment and accountability that

motivated my interest in this study, I made an effort to bracket my own opinions when critically reflecting on the data informing this dissertation work. As a White, middle-class woman, I intentionally drew from a polyphony of voices and perspectives to not center the perspectives or thoughts of any one group as representative of the whole.

As a constructivist, my perspective is that all meaning is socially constructed, and there is no way to eliminate my own perspectives from my research; however, I practiced reflexivity and actively searched for disconfirming evidence throughout the research process in order to challenge my own thinking and perspectives about the topic.

Critical Policy Analysis Framework

This three-manuscript dissertation investigates Virginia assessment and accountability policies by applying critical frameworks to policy analysis (Diem et al., 2014; Diem & Young, 2015; Young, 1999; Young & Diem, 2014, 2017, 2018). Across the three papers, a variety of critical frameworks will be applied to assessment and accountability policy to critique and promote enlightened action and change by policymakers around assessment and accountability policies more holistically.

Paper One: A Critical Policy Analysis of Virginia State Assessment and Accountability Policies

In paper one, I used Critical Policy Analysis (CPA; Young & Diem, 2017) informed by Rein's value-critical approach to policy (1976) and Guba's (1984) definitions of policy (see Table 1), to analyze current assessment and accountability policies in the state of Virginia. CPA's tenets require dissecting histories and origins of policies, differences between policy rhetoric and experience, power dynamics by considering policy winners and losers, social stratification within the policy by examining who is privileged and oppressed, and groups who

resist oppressive policies and systems (Young & Diem, 2017, p. 4). Varying marginalized identities including race, class, and ability status were considered through this analysis, and Bradbury's (2020) framework was also used to look at race within a critical policy analysis of assessment and accountability policy.

The purpose of this study was to explore each tenet of CPA applied to assessment and accountability policies in Virginia. First, the political history that led to the current assessment and accountability policies was investigated and detailed in narrative form. This included a detailed review of literature around both the history of assessment as well as the political policy landscapes that led to the current federally required standardized testing in Virginia. Included in this narrative are references to state-wide attempts to conceptualize and implement alternative assessments as well as the successes and pitfalls of those plans.

Next, the rhetoric of Virginia standards, assessment, and accountability policies was investigated, with special attention to power and tendencies toward social stratification. These policies were assessed by examining the written policy language and observations of interactions of state-level policymakers at Virginia Board of Education (BOE) work sessions and meetings. Via a media review, groups and individuals who have resisted the standards, assessment, and accountability systems were highlighted. Additionally, educators who have experienced the ambiguity in state policy, with the policy-in-action diverging from the policy-in-intention were considered.

Paper Two: The Standards of Learning and the 5 C's: Virginia State Policies in Action

In paper two, I examined the lived experiences of teachers who have implemented the multiple and ambiguous state policies in one school district. Flyvbjerg's (2001) approach to phronetic research, which infuses interpretivism with an analysis of power through case study

methods was utilized to look at contrasts between rhetoric and practice. To help conceptualize the differences between rhetoric and practice, Guba's (1984) definitions of policy, in which he sorted eight definitions of policy into three created categories: policy-in-intention, policy-in-action, and policy-in-experience were applied (see Table 1). Within each of these policy definition types, there were multiple variations and definitions of a single policy, except for policy-in-experience, which included only the single definition of how the policy is understood by those who encounter it. For a simplified explanation of the definitions within each type, including the level of interaction with the policy and what the policy looks like at that level, see Table 1.

This study highlighted, from the perspectives of Virginia policymakers and teachers who were interviewed, differences between policy-in-intention and policy-in-implementation around the high-stakes Standards of Learning standardized tests and the introduction of performance assessments and higher order skills with the Profile of a Virginia Graduates, including the 5 C's: critical thinking, creative thinking, collaboration, communication, and citizenship. In the context of education where policymakers are far removed from teachers in classrooms who are implementing the policies around assessment and accountability, this framework was a helpful guiding concept. This case was an exemplar of the effects of ambiguity in policy intentions that produced differences in rhetoric and practice explored in paper one.

Table 1

Guba's (1984) Definitions of Policy

Guba's Typology	Policy defined as	Interaction with Policy	Looks like
Policy-in-intention	-Goals -Regulations -Guides for decision-makers -Problem-solving	Distant/Writer/ High-level administrator	-Ends -Rules -Guidelines -Tactics
policy-in-action	-Rules for policy enforcers -Conduct norms -Output of the system	Close/Implementer	-Expectations -Norms -Effects
policy-in-experience	-Impact of the policy	Direct Experience	-Experiences

Note. Adapted from Figure 1 (Guba, 1984, p.65)

Paper Three: What Does Success Look Like? Teachers' Realities and Reimaginings in K-12 Schools

In paper three, written with Dr. Helen Min, Critical Systems Heuristics (CSH; Ulrich, 1996) were applied to help consider not only the effects of the assessment and accountability system as-is, but also how the educational system *ought* to be from teachers' perspectives within this context. The CSH framework helps to illuminate the entanglement of a system and how problems within a system are not easily disentangled and solved, and how solutions that fix particular problems often create others (Ackoff, 1974; Ulrich, 1996). In thinking about educational policy, this is helpful as a tool to systematically look at how to connect teacher agency and political structure as advocated for by Flyvbjerg (2001). In this paper, the power/control components of CSH were central to analysis. For more information about the components of CSH, see Table 2.

Table 2

Critical Systems Heuristics Adapted to a System of Education

	Roles/ Stakeholders	Role Specific Concerns/Stakes	Problems/ Stakeholding Issues	
Values & Motivations for Improvement	*Beneficiary* Who is/ought to be the beneficiary of the education system?	*Purpose* What is/ought to be the purpose of the education system?	*Measure of Improvement* What are/ought to be the measures of success for the education system?	
Power/ Control	*Decision Makers* Who is/ ought to be in control of the conditions of success in the education system?	*Resources and Attributions* What conditions of success are/ought to be under control of the education system?	*Environment* What conditions of success are/ought to be outside the control of the decision makers?	Teachers (The Involved)
Knowledge	*Planner/Expert* Who is/ought to be providing relevant knowledge and skills for the education system?	*Expertise* What are/ought to be the relevant new knowledge and skills for the education system?	*Guarantee* What are /ought to be the assurances of successful implementation?	
Consequence/ Legitimacy	*Witness* Who is/ought to be representing interests of students in the education system?	*Emancipation* What are/ought to be opportunities for students to have expression and freedom from the constraints of the educational system?	*Worldview* What space is/ought to be available for reconciling different worldviews regarding the education system among teachers and students?	Students (The Affected)

Note. Adapted from Figure 3 (Ulrich , 1996, p. 44)

In this study's application of CSH, teacher perceptions of student and teacher success within the current system of assessment and accountability are explored. Additionally, teachers' views on a reimagined system that could promote better success for students and teachers were discussed. This paper extends the other two papers by presenting an action-orientation related to the critiques of papers one and two. Thus, it indicates the direction teachers would like the education system to move toward and what structural barriers are preventing those changes. One of the barriers identified and discussed is the assessment and accountability policy system itself,

which teachers identified as a barrier to reimagining the ways they want to teach and measure success. This paper illuminated the putative ends of quality education over a myopic view of the means framed by standards, assessment, and accountability policy.

For policymakers considering standards, assessment, and/or accountability policy change, the teacher perspectives of the barriers to change reflected in this study are informative. This paper comports with the "educated hope [that] we move forward not toward 'new' dystopias but toward the anticipatory not-yets of utopias-to-come" (Berbary & Mohamed, 2022, p.6). While critics of utopias have long considered this line of thinking "either politically immature or morally dangerous," a critical perspective on reimaginings, which considers deeply the historical origins of the system as it is, can lead to "transformative possibilities with an awareness of its limitations" that "can invigorate contemporary political thinking" (Zamalin, 2019, pp. 1-2). Kelley (2022) reminds us that "without new visions we don't know what to build, only what to knock down" (p.xii). So, inspired by the freedom dreaming of Love (2017, 2019, 2023) and Kelley (2022), and in line with CSH, in this paper we asserted that in order to enact fruitful change, systems first need to be reimagined, and that policy is a prohibitive barrier.

The themes within this study were formed from teacher interviews. A version of this paper was written with second author Dr. Helen H. Min and has been accepted as a peer-reviewed chapter in the book, *Reimagining Systems Thinking in a Post Pandemic World*, edited by Azukas and Kim (Forthcoming). To more closely align with the themes of this dissertation, a slight reframing of this study was written in order to more directly highlight the role of assessment within the education system and its role in teachers' visions of change.

Significance

Together, the critical perspectives from these three papers help illuminate the differences between policy in rhetoric/intention, implementation, and experience and illuminate the winners and losers from standards, assessment, and accountability policy in order to determine if the Virginia assessment and accountability system actually promotes the desired knowledge and skills asserted in the Profile of a Virginia Graduate for students regardless of race, class, or ability identity. They exemplify the challenges of federal and state policy implementation with multiple embedded ambiguous goals, and how educators at the school level are interpreting and acting on those policies. For considerations of implications, recommendations, and future research to consider following these three manuscripts as a whole, see the epilogue.

This study contributes to the scholarly works surrounding accountability and assessment systems in the current era by illuminating, from critical perspectives, the problematic issues baked into educational policy surrounding K-12 accountability and assessment policy in Virginia. Additionally, places where the goals and outcomes of the policies are misaligned as well as mis-interpreted are exposed. We know that at a national level, critical scholars have been questioning if standardized assessment and accountability policies are equitable and helpful for the authentic engagement and success of all students. This dissertation highlights the specific implications of these dilemmas within the Virginia assessment and accountability policy landscape. This research is significant because it clarifies where the putative intentions of assessment policy may be misguided and illustrates how insidious forms of oppression are embedded within the policies' rationalized maintenance of the status quo.

This is especially important as accountability and assessment policies in Virginia are currently under review and could potentially change. Virginia policymakers are at a crossroads

with standards and assessment policy, caught between the status quo of standardized testing, reliant on discrete content knowledge standards, and the potential for performance tasks to test authentic 21st century learning outcomes aligned with the Profile of a Virginia Graduate. They have created competing policies and are sending mixed messages to district administrators and teachers. As decisions are being made whether or not to further eliminate certain subject matter and grade-level standardized tests, teachers and policymakers should be aware of how these policy changes may or may not promote the narrowing of the curriculum and how that in turn impacts student learning. The study illuminates what can occur when policymakers merely add policy mandates that potentially conflict with the dominant standards-based approaches. For instance, what happens when policymakers add the Profile of a Virginia Graduate with a mandate to teach the 5 C's: critical thinking, creative thinking, citizenship, communication, and collaboration, onto the current system dominated by the Standards of Learning (SOL) and high-stakes testing? What happens when state policy makers add conflicting and ambiguous mandates onto teachers' plates?

If certain untested concepts are ignored, it can lead to both difficulties in higher-level courses as students progress through their education as well as the graduation of under-prepared citizens. The results of the study illustrate to policymakers the inherent divergence in fundamental goals behind two conflicting assessment and accountability approaches. It also illustrates how assessment policy is significantly influenced by partisan politics despite the technical-rational veneer. The results illustrate the challenges and choices faced by policymakers who are considering reforms to policy systems with goals and values that vary from traditional efficiency-oriented systems. Lessons learned from trying to force the Virginia standards and accountability-based SOL approach simultaneously with the Profile of a Virginia Graduate low-

stakes alternative performance approach lends insight to future policy decisions. As Virginia K-12 classrooms continue to deal with the ramifications of the pandemic and changes in policy at the state level, this critical analysis of assessment and accountability systems provides relevant information to teachers, activists, policymakers, and families who want to advocate for systems that benefit all Virginia students.

References

Ackoff, R. L. (1974). *Redesigning the future: A systems approach to societal problems.* John Wiley & Sons.

Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy*, *30*(1), 39–62.

Aurora Institute, The Center for Assessment, The Center for Innovation in Education, Envision Learning Partners, The Great Schools Partnership, and KnowledgeWorks (2021). *Measuring forward: Emerging trends in K-12 assessment innovation*. https://knowledgeworks.org/resources/emerging-trends-k12-assessment-innovation/

Azukas, E. & Kim, M.(Eds). (Forthcoming) *Reimagining Systems Thinking in a Post Pandemic World*. IGI Global.

Berbary, L. A., & Mohamed, L. A. (2022). Longing as method: A rant on yearnings for our world, academia, and utopian futurities beyond liberalism(s). *Qualitative Inquiry*, 10778004221111379.

Berman, P., & McLaughlin, M. W. (1976). Implementation of educational innovation. *The Educational Forum, 40*(3), 345-370.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas Accountability System. *American Educational Research Journal*, *42*(2), 231–268.

Bradbury, A. (2020). A critical race theory framework for education policy analysis: The case of bilingual learners and assessment policy in England. *Race, Ethnicity and Education*, *23*(2), 241-260.

Culpepper-Ofori, D. (2012). *The development of tracking and its historical impact on minority students (Publication No.* 3509468) [Doctoral dissertation, Walden University]. ProQuest Dissertations Publishing.

Diem, S., Young, M. D., Welton, A. D., Mansfield, K. C., & Lee, P. L. (2014). The intellectual landscape of critical policy analysis. *International Journal of Qualitative Studies in Education, 27*(9), 1068-1090.

Diem, S., & Young, M. D. (2015). Considering critical turns in research on educational leadership and policy. *International Journal of Educational Management*, *29*(7), 839-850.

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press.

Gould, S. J. (1996). *The mismeasure of man: Revised and expanded edition.* W. W. Norton & Company.

Guba, E. G. (1984). The effect of definitions of "policy" on the nature and outcomes of policy analysis. *Educational Leadership*, *42*(2), 63-70.

Heinecke, W. F., Curry-Corcoran, D. E., & Moon, T. R. (2003). US schools and the new standards and accountability initiative. *Educational leadership in an age of accountability: The Virginia experience*, 7-35.

Henderson, M. B., Houston, D. M., Peterson, P. E., & West, M. R. (2022). Hunger for stability quells appetite for change: Results of the 2021 Education Next Survey of Public Opinion. *Education Next*, *22*(1), 8-26.

Hollingworth, L. (2009). Unintended Educational and Social Consequences of the No Child Left Behind Act. *The Journal of Gender, Race, and Justice*, *12*(2), 311.

Kelley, R. D. (2022). *Freedom dreams: The black radical imagination.* Beacon Press.

Labaree, D. F. (1997). Public goods, private goods: The American struggle over educational goals. *American Educational Research Journal*, *34*(1), 39-81.

Love, B. L. (2017). "She has a real connection with them": Reimagining and expanding our definitions of Black masculinity and mentoring in education through female masculinity. *Journal of Lesbian Studies*, *21*(4), 443-452.

Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom*. Beacon Press.

Love, B.L. (2023, March 2). *Abolitionist teaching for educational justice*. Walter N. Ridley Lecture, University of Virginia, Charlottesville, VA.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.

Mandery, E. (2022). *Poison ivy: How elite colleges divide us*. The New Press.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. (ED 226006). ERIC. https://files.eric.ed.gov/fulltext/ED226006.pdf

Newmann, F. M., King, M. B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review, 67*, 41–74.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Harvard Education Press.

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002). https://www.govinfo.gov/content/pkg/PLAW-107publ110/pdf/PLAW-107publ110.pdf

Odden, A. (1991). *Education policy implementation*. State University of New York Press.

Kendi, I. X. (2016). Why the academic achievement gap is a racist idea. *Black Perspectives.*

Randall, J. (2021). "Color-neutral" is not a thing: Redefining construct definition and

representation through a justice-oriented critical antiracist lens. *Educational*

Measurement: Issues and Practice, *40*(4), 82-90.

Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting White Supremacy in

assessment: Toward a justice-oriented, antiracist validity framework. *Educational*

Assessment, 1-9.

Rein, M. (1976). *Social science and public policy.* Penguin Books.

Smith, M. L., Edelsky, C., Draper, K., Rottenburg, C., & Cherland, M. (1990). The role of

testing in elementary schools (CSE Tech. Rep. No. 321). Los Angeles: University of

California. *Center for Research on Evaluation, Standards, and Student Testing*.

Ulrich, W. (1996). A primer to Critical Systems Heuristics for action researchers. *Independent*

Research Monograph, Centre for Systems Studies, University of Hull.

Viennet, R., & Pont, B. (2017). Education policy implementation: A literature review and

proposed framework. *(OECD Education Working Papers No. 162)*.

https://doi.org/10.1787/fc467a64-en.

Wilson, M. (Ed.). (2004). *Towards coherence between classroom assessment and accountability*.

University of Chicago Press.

Yilmaz, M., Buchanan, M. B., & Martin, J. (2023). Disrupting majoritarian narratives in high

stakes testing: A critical race theory analysis of assessment. *Equity in Education &*

Society, *2*(1), 28-42.

Young, M. D. (1999). Multifocal educational policy research: Toward a method for enhancing

traditional educational policy studies. *American Educational Research Journal*, *36*(4),

677-714.

Young, M. D., & Diem, S. (2014). Putting critical theoretical perspectives to work in educational

    policy. *International Journal of Qualitative Studies in Education, 27*(9), 1063-1067.

Young, M. D., & Diem, S. (2017). Introduction: Critical approaches to education policy analysis.

    In M.D. Young & S. Diem (Eds). *Critical approaches to education policy analysis:*

    Moving beyond tradition, 1-13.

Young, M. D., & Diem, S. (2018). Doing critical policy analysis in education research: An

    emerging paradigm. *Complementary research methods for educational leadership and*

    policy studies, (pp. 79-98). Pagrave Macmillan.

Zamalin, A. (2019). *Black utopia: The history of an idea from Black nationalism to Afrofuturism*.

    Columbia University Press.

Paper One: A Critical Policy Analysis of Virginia State Assessment and Accountability Policies

This paper will present the results of a review of the literature of the standards, assessment, and accountability movements in the U.S. and Virginia with the ultimate goal of framing the alternative assessment standards and assessment policy debate currently under discussion in the State of Virginia and providing the context for the second part of the paper that presents a critical policy analysis (CPA) of Virginia's current assessment policy. This paper sets the stage for understanding the experiences of educators implementing those policies that are examined in papers two and three.

While education policies undergo periodic review, over the last several years, debates around best practices in assessment and accountability have drawn significant attention, exacerbated by the COVID-19 pandemic that disrupted the status quo systems. Within this context, state policymakers have a renewed interest in critically examining their current standards, assessment, and accountability systems while considering innovations for change. After reviewing the context of national and state accountability and assessment policy, this paper provides the results of a critical policy analysis of current Virginia state policy. Young and Diem (2017) offer critical policy analysis (CPA) as a means to study and critique existing polices as well as a pathway to suggest alternative options. In order to understand differences between rhetoric and practice around assessment and accountability policy in Virginia, a deeper dive into the language used is needed. CPA's tenets require dissecting histories and origins of policies, differences between policy rhetoric and experience, examining where power is at play and who is winning and losing with the enactment of the policy, social stratification within the policy and who is being privileged and oppressed, and groups who resist oppressive policies and systems

(Young & Diem, 2017, p. 4). The research questions for this paper are directly derived from the tenets of CPA. With the data collected, the following questions will be investigated:

1. What is the policy and political history of assessment and accountability policy in Virginia?

2. What is the rhetoric and values surrounding accountability and assessment policy in the state of Virginia?

   a. Where does mixed policy messaging exist?

   b. How might the conflicting values and intentions embedded in current assessment policy create ambiguity for educational practitioners?

3. Where are notions of resistance occurring that indicate divergences in policy from practice to subvert oppressive systems?

Definitions

As is true in most fields, in the accountability and assessment literature, there are families of terms that refer to similar practices, but with nuanced differences. In order to clarify the language, it helps to begin by defining relevant terms that will be oft referred to.

Assessments can be at the classroom, state, national, and international level and can be given in a variety of formats. Of importance to this work is the difference between *standardized* and *performance-based* assessments. Standardized assessments are often multiple-choice in format, but sometimes can use technology to allow the questions to be given in a number of formats. However, within these assessments there is generally only one correct answer for each question that is sought as an indication of mastery. Performance-based assessments, also sometimes referred to as performance assessments, project-based assessments, or alternative assessments, instead, require the student to answer a question or solve a problem in a way that

there are likely many correct answers. These types of assessments are often scored using rubrics rather than simply as right and wrong. The emphasis here includes not just the answers, but how students arrived at them.

Each of these types of assessments, and sometimes a combination of them, are given within an *accountability system*. An accountability system includes both the assessment and the stakes assigned to the results. Accountability can be defined in a myriad of ways, including, on one end, supporting schools in meeting their goals to, on the other, a metric for control (Heinecke et al., 2003). Accountability can be at the student-, teacher-, school-, district-, state-, and even federal-level. These levels of accountability can be grouped by thinking of individual, institutional, professional, and political accountability (Beadie, 2004, p.35). Additionally, the stakes can either be high or low. High-stakes accountability at the student-level often includes barriers to graduation or grade promotion for failure to achieve a certain benchmark on an end-of-year assessment. For teachers, high-stakes professional accountability can refer to credentialing that is required to become a teacher as well as having pay and/or job security tied to students' testing performance. High-stakes institutional accountability can be used to compare and punish schools. Schools are often put on a distribution, compared publicly in the media, and penalized for low aggregate assessment scores. When aggregate scores are low, the agency that teachers and school leaders have is often minimized with more top-down management through higher surveillance and required implementation of certain curricula, with the state controlling most of the decision-making within those schools. A low-stakes assessment, in contrast, is used as a point of data collection, but is not usually used as a negative incentive or punishment if performance is not high. Both standardized and performance-based assessments can be low- or high-stakes, depending on how the data is used once collected. In some instances, researchers

have found that performance assessment systems have been interpreted and used in the same manner as high stakes standardized testing formats (Smith et al., 1994).

When considering how these various assessment and accountability systems are enacted, it is relevant to note some paradigmatic assumptions behind different models of learning upon which assessment systems are built. If learning is viewed through a constructivist lens, students are acknowledged as adding new information into the schemata of what they already know. This perspective holds a level of comfortability with differences in learners and Gardner's (1983) multiple intelligences (Mabry, 2004; Smith et al., 1994). An opposing view categorizes learning in terms of behaviorism, looking at added knowledge as a changed behavior which can be quantified, the dominant view behind more standardized assessment approaches.

These different perspectives are tied to different theories of action about how to hold students, teachers, and school systems accountable. Systems that are committed to high-stakes standardized testing operate under the assumption and theory of action that incentives and punishments will change the behavior of both teachers and students to make them work harder and therefore achieve more (McDonnell, 2005; Newmann et al., 1997; Nichols & Berliner, 2007). A conflicting theory of action contends that authentic alternative assessments, along with professional educators who have the correct preparation and support, are more likely to improve instruction and learning (See paper three; Stosich et al., 2018). Added to this theory of action is the motivation that "what is assessed is what gets taught" (Smith et al. 1994, p. 4). Therefore, the perception of the policy by the teachers matters, as the policy-in-action can change from low-stakes to high-stakes depending on how the data itself is handled (Smith et al., 1994).

Conceptual Framework

Critical Policy Analysis

Young and Diem (2017) advocate that CPA is both conceptual framework and methodology (p 2). Therefore, the tenets of CPA are foundational to the methods of this paper. Critical education policy studies aim to look simultaneously at socialization through schools and society, center values in analysis and reject the notion that analysis can be objective, acknowledge that power structures within schools are often structured by policy, and ask "whose values have been validated" (Prunty, 1985, pp. 135-136). CPA is derived from a paradigm of critical theory with the ontological assumptions associated with historical realism, in which it is understood that the cultural norms of society get reified over time to construct a reality that can be hard to change (Lincoln et al., 2018). Epistemologically, this means that understandings of reality are subjective, and that we are all always seeing the world through our own values and lenses. This informs methodological assumptions that require that data be collected from multiple perspectives in order to better understand how individuals are understanding and interpreting realities. In critical perspectives, the voices of those who have been historically marginalized are often centered with an aim toward critique and/or activism toward change.

Rein (1976), Guba (1984), and Young and Diem (2017) offer overlapping and convergent approaches that will inform this CPA. Rein's value-critical approach to policy analysis asks that analysts recognize: (a) that while policies may remain unchanged, societal goals and purposes for them may change; (b) that the practice associated with policies must be considered; (c) that historical perspectives illuminate recurring problems and context; (d) the need to distrust orthodoxy, always questioning the status quo and looking for places where alternative solutions would better meet the needs of intended policy audiences; (e) that there is a role for the policy

analyst as a moral critic, questioning the morals at play within an established policy and which considerations are ignored with the centering of others; and (f) that policy and policy research are used within the political process, sometimes in unintended ways.

There is clear overlap here with Young and Diem's (2017) CPA guidelines to (a) pay attention to how policies came to be; (b) dissect differences between policy rhetoric and outcomes; (c) unearth notions of power; (d) consider how policy intentionally or unintentionally produces social stratification; and (e) amplify the voices of people who are resisting oppressive systems. Guba's (1984) definitions of policy: policy-in-intention, policy-in-action, and policy-in-experience (as defined in Table 1 in the Linking Document) help to provide a detailed framework about the varying concepts meant by the word *policy* and how to interrogate those differences. These definitions will be used to help clarify the levels and interpretations of policy within the results. Lastly, Bradbury's (2020) framework that specifically attends to issues of race within educational CPA will be used supplementally. This framework encourages the researcher to ask questions such as "Is the creation of the policy problem a distraction from matters of racial equity?" and "How does this policy present those in power as caring about racial equity (even where this is not the result)?" (Bradbury, 2020, p. 247). This will ensure that race is specifically addressed in this application of the CPA tenets outlined by Young and Diem.

Politics as a Component of CPA

The idea of political spectacle, rooted in Edelman's theory of symbolic politics, will be applied to educational policy to investigate when policymakers care more about "the appearance of having done something" as opposed to the actual effects of the policy itself (Smith et al., 2004, p. vii). Political spectacle can be seen by interrogating symbolic language, pitting political leaders in roles of champions and enemies, and disconnections between calls on rationality and

the ends of policies themselves. As it turns out, "Assessment policy epitomizes the political spectacle" (Smith et al., 2004, p.46), and therefore is a useful framing in the discussion around rhetoric versus outcomes. This aspect of the conceptual framework addresses the partisan political influence of standards, assessment, and accountability policies.

Stein's (2004) culture of education policy will also be used as a framework through which to analyze political spectacle within standards, assessment, and accountability policies. Stein asserts, in her analysis of Title I policy from the federal level to the classroom level, that the language of education policy was framed from a deficit-based perspective in which the students who are in most need of educational policies for their improvement are labeled as deviant from the norm. She explains that this framing of children and schools as "inadequate" and "failing" by the policy "encourages practitioners to talk about [students and entire schools] in terms of their inadequacy, deficiency, and deprivation" (Stein, 2004, p.ix). She acknowledges how policymakers create a policy culture that undermines putative policy beneficiaries. In order to understand policy, both the language and behaviors of the policymakers are an important component of CPA, including their political motivations. In order to use this framework, the researcher focuses on the culture of the policy process itself as problematic rather than a focus the on those presumed to benefit from the policy. This framework raises the question if government agencies and policies can really serve as an ameliorative force or if policy language continues to exacerbate inequity.

Lastly, divergent beliefs about the purpose of education will be used to help uncover the tensions between the political perspectives of policymakers. While intersecting with Rein's (1976) values-centered policy analysis, this lens identifies the particular values and purposes central in the American context. Labaree (1997) identifies three competing archetypes for the

purpose of school: preparing democratic citizens, training economic laborers, or allowing for economic mobility. In this conception, the idea of preparing citizens for love of learning or personal development is absent. Marshall and colleagues (1989) frame the discussion through the lens of sometimes competing American values of equity, quality, choice, and efficiency. These values have been in tension with each other throughout the history of U.S. education policy and help explain the motivations behind the politically-motivated actions of policy actors.

Review of the Literature: Political and Policy History of Assessment and Accountability

In order to have a thorough understanding of the roots and origins of the policy system central to this CPA, the first part of this paper will narrate a history of assessment and accountability. In the early 1900s, trends of industrialization were applied to people, with a desire for them to be measured and classified. Many assessment systems, including standardized testing, are firmly connected to the scientific study of eugenics (Au, 2016; Gould, 1996). While the originator of the IQ test, Alfred Binet, asserted that intelligence was not hereditary and could not be quantified precisely by a scale in the early 1900s, by the 1920s, eugenicists adapted Binet's work to be used in the U.S. and perpetuated the narrative that whiter, wealthier people were genetically superior and more intelligent than their counterparts (Au, 2016; Berliner & Biddle, 1997; Culpepper-Ofori, 2012; Gould, 1996). The work of these men had lasting impacts on American societal perceptions of people and the direction of assessment in the United States.

Cases of cultural bias in testing did not disappear with the decline of eugenics. In the late 1970s, a case in California, *Larry P. v Riles* led to the ban of IQ tests in determining the need for special education services for Black students due to its biases that led to their over-identification (Prasse & Reschly, 1986). These tests, banned in California, are still actively used in other states across the U.S. today.

Early Accountability Movements

Accountability has long been considered in education but in ways different from the concepts of accountability today. From the 1920s to 1940s, efficiency and quality were the central values, with accountability monitored at a more local level (Heinecke et al., 2003). In the 1950s and 1960s, equity was added as a value, and federal accountability was considered with the implementation of the Elementary and Secondary Education Act of 1965 (ESEA). With its passage, federal requirements for standardized testing began.

Standardized Testing for School Accountability

ESEA (1965) required norm-referenced standardized tests in order for schools to receive particular federal funds; however, these test results were not available for public release (Heinecke et al., 2003). In the early 1970s, the National Assessment of Educational Progress (NAEP) was instituted as a means to compare student achievement data, but as it used only a representative sample, it was not intended for individual student accountability.

In 1981, the U.S. Secretary of Education created the National Commission on Excellence in Education (NCEE) in order to address "the widespread public perception that something is seriously remiss in our educational system" (NCEE, 1983, p. 6). The resulting report, *A Nation at Risk*, took an alarmist tone as it presented the ways that American schools and students had fallen behind those in other nations. It was from this report that a federal call for standardization was put into action. This set of waves of federal policy movements that were pushed across the states including standards-based reforms and the creation of highly detailed accountability policies with assessment policies embedded within them.

The basis of this introduction of standardized tests was the claim that achievement was declining in the United States, especially in comparison to other countries (NCEE, 1983).

Interestingly, when comparing data collected at the time, only SATs showed this decline. Since the SATs are taken voluntarily by students who are considering higher education, the non-representative sample of students' aggregate scores could indicate as easily an expansion of the pursuit of post-secondary education as a decline in achievement (Berliner & Biddle, 1997). The PSAT (Preliminary SAT) and NAEP data collected at the same time supported this theory, as they both included a representative sample of students. PSAT scores did *not* decline and NAEP scores actually increased. In *The Manufactured Crisis*, Berliner and Biddle (1997) ask why public schools could be considered to be in crisis if the evidence supported the contrary. Their research-backed response was that the manufactured crisis was largely a political move to use U.S. public education as a scapegoat for larger issues related to unions and for the promotion of further privatization and marketization of schools and other social institutions. This is an illustrative example of both political spectacle in education (Smith et al., 2004) and Rein's (1976) assertion around manipulation of data for partisan policy goals. Standards, assessment, and accountability policy became a vehicle for these political battles.

As part of ESEA, the Clinton administration championed *Goals 2000*: Educate America Act in 1994, which incentivized states through grants to use standards-based reform, including testing in math and reading (Schwartz et al., 2000). These national calls for reform were filtered through states, who are responsible for their own systems of education (Heinecke et al., 2003). According to Grissmer and Flanagan (1998), in the early- to mid- 1990s, two states, North Carolina and Texas, implemented high-stakes testing and saw larger growth on the NAEP than any other states. They concluded "the most plausible explanation for the test score gains are found in the policy environment," which they defined as the state creating standards, curriculum, and assessments and holding schools accountable (Grissmer & Flanagan, 1998, p. i). However,

they also acknowledged concerns about external validity for their own study, noting that their findings may not be generalizable. Despite this warning, many other states in the U.S. formulated and adopted high-stakes standardized tests as a form of accountability for students and schools prior to the federal mandates of the No Child Left Behind Act in 2002 (NCLB; Dee & Jacob, 2011). As with most decisions in today's neoliberal society, there are political and economic incentives for this. Standardized assessments allow for the ranking and comparison of schools which can be used to attract large corporations and bolster, hypothetically, more profitable competition.

While many stakeholders were concerned about the idea of external accountability promoting school-to-school comparisons, an assessment study in the mid-1990s supported that internal accountability was more important for authentic student learning (Newmann et al., 1997). They stated, "external accountability alone offers no assurance…for student learning…. It is also possible that strong accountability can be achieved within a school community, without prescriptive mandates from a district or state" (Newmann et al., 1997, p. 62). These results support that a less centralized approach to accountability and assessment could be effective.

In Nichols & Berliner's (2007) interpretation of Campbell's law they state: "You can have (a) higher stakes and less certainty about the validity of the assessment or (b) lower stakes and greater certainty about validity. But you are not likely to have both high stakes and high validity. Uncertainty about the meaning of test scores increases as the stakes attached to them become more severe" (p. 27). Others agree that the entire system of high-stakes accountability is in opposition to validity, with many teachers' "test preparation" under such stressful circumstances being tantamount to outright cheating (Cannell, 1989; Chen, 2021; Smith, 1991; Smith & Fey, 2000). Regardless, as media and policies have portrayed quantitative standardized

tests results as objective, the public largely accepts them at face-value as valid, and their use became codified into law in 2002 with the passage of the NCLB Act (2002).

Adoption of No Child Left Behind (NCLB)

While many states had been implementing high-stakes accountability and standardized assessment systems in the 1990s, once the NCLB act (2002) was adopted in the early-2000s, almost all states shifted their accountability system and employed large-scale standardized testing. This was an expansion of the requirements of ESEA, as it mandated standardized testing within an accountability system that required schools to make yearly progress toward higher overall levels of proficiency on math and reading tests. Nebraska was the only state that did not immediately employ a high-stakes testing system, choosing instead to utilize the School-based, Teacher-led Assessment & Reporting System (STARS), which employed several local assessments to ensure that state standards were being met (Ruff, 2019). The state board reviewed assessments to ensure alignment, but this low-stakes approach allowed school contexts to adapt testing practices that best met their students' needs. Nebraska was granted a federal waiver for STARS, and the system gained strong support across the state due to its success. The program was in place until its key organizer at the state-level retired in 2008. Shortly thereafter, the state moved to a system that was like that of the rest of the U.S., another example of the importance of political actors in policy success (Ruff, 2019).

Critical Concerns about Alignment

The NCLB era raised a host of issues related to race and equity in assessment policy. Often, especially in popular media, gains shown on assessments in the early 2000s credit NCLB with closing achievement gaps and improving math scores (Leonhardt, 2022). However, in a study of Chicago Public Schools that utilized longitudinal student-level data and compared

district-level data to other large cities in the Midwest, evidence supported that student gains shown in high-stakes testing were not present in other low-stakes tests on the same subject, suggesting that what looked like jumps in achievement due to the high-stakes accountability system were actually likely increases in test taking skills (Jacob, 2005). Even if we were to credit NCLB policies with increased scores, the growth has not sustained. Of special importance are concerns about how classism and racism skew results. Finally, the ways that assessments and educational goals are often misaligned adds to the arguments against standardized testing as an effective educational policy.

Socioeconomic Concerns

There is evidence that low-income students suffer under a high-stakes accountability system. For example, a pre-NCLB study found that 81% of variation in standardized test scores could be explained by variations in the school's percentage of low-income students (Rusk, 1998). Additionally, an analysis on NAEP data found that although a simple analysis may show some growth, these results did not hold true when the analysis controlled for family income (Marchant et al., 2006). Berliner & Biddle (1997) refer to these correlations as the ability of wealthy parents to "'buy' intelligence-test points for their children," which is largely contributed to by the inequities in school funding (p. 49).

Racial Concerns

There were additional concerns about the impact of high-stakes testing on students of color raised during this period. If we consider the historical foundations of standardized assessments presented earlier, it is difficult to claim that the standardized practices of today could be for equity when the very formulation of the assessment practice was created with the opposite goal in mind. As Chira (1992) stated, "Widely used standardized tests may be

hampering efforts to improve math and science education, and they hurt minority students the most" (as cited in Berliner & Biddle, 1997, p. 318).

NCLB proponents touted the policy's putative impact on closing the *achievement gap,* a term now often considered racist itself, as one of its main goals; however, a growing body of literature maintains that the system actually did the opposite (Au, 2016; Gould, 1996; Hollingworth, 2009; Inoue, 2015; Love, 2019; Randall, 2021, Randall et al., 2022; Yilmaz et al., 2022). These scholars challenge the taken-for-granted nature of standardized assessments as objective measurements. One of the issues with this assumption of objectivity is that it puts the onus of blame on students who have fallen behind rather than the tests or accountability system itself. Scholars critical of high stakes accountability and standardized testing reject this assumption, and the claim that these assessments promote equity.

Through their own counter-narratives and research on their perspectives of such accountability systems, teachers and administrators have challenged district and policymaker perspectives on the benefits of this type of testing. According to Yilmaz and colleagues (2022) "All of them seemed to agree that the high stakes nature of the tests are not serving students but the district-level staff seemed more open to the idea that standardized assessments have a place and can serve a productive purpose" (p.8). This illustrates the role of administrators in maintaining the authority of state-level policy in implementation at the school and classroom levels. Students' counter-narratives about standardized testing also raise concern about the accountability system. For instance, in a study of African-American fourth graders, students shared their beliefs that education was largely test preparation and emphasized the anxiety they felt toward the end-of-year tests (Wasserburg, 2017). These types of studies are ubiquitous during this era. The students explicitly spoke to the narrowing of the curriculum stating, "We

don't do social studies anymore because our teachers, they want to do the basic ones that they know that we're gonna get on the [Florida Comprehensive Assessment Test], like math and reading" (Wasserburg, 2017, p.44). This leads to additional concerns about the misalignment of desired learning outcomes and assessment and accountability policies.

Concerns around Misalignment between Education and Assessment

The pressures of high-stakes testing are also reported to result in a narrowing of the curriculum (Smith et al., 1990; Tracey, 2005). In many research studies and media reports, teachers have corroborated Wasserburg's (2017) findings indicating that due to the pressures put in place by NCLB, they ignored portions of the curriculum, including both a deeper engagement with tested subjects and subjects that were not tested, to spend more time on concepts likely to appear on the high-stakes tests, even when they found other concepts more important to the lived experiences of their students (See paper two; Smith et al., 1990; Tracey, 2005). As Berliner & Biddle (1997) state after advocating for a new curriculum that is more relevant to 21st century learners: "It is difficult to have both a new curriculum and traditional, multiple-choice tests. The new curricula and traditional tests make poor bedfellows" (p. 318). This ambiguity is central to the results of this study.

As students enter both higher education and the workforce, many researchers have uncovered a gap between the expectations of professors and employers and students' skills. Studies support that mathematics skills that professors consider the most important are among those that students struggle with the most (Corbishley & Truxaw, 2010; Er, 2017). Despite mathematics being one of the two subjects most focused upon in systems of standardized testing and accountability, the necessary skills are not translating into deep learning or long-term applications. Additionally, employers have repeatedly noted that while soft-skills, such as

problem-solving and communication, are among the most important skills they look for in employees, they feel that new hires are lacking in these skills (National Association of Colleges and Employers, 2019; Stewart et al., 2016). These are the unintended consequences of high stakes accountability and assessment systems implemented since the mid-1980s.

Implementing Alternative Systems of Accountability

With these arguments in mind, in the 1990s, academics, research centers, and policymakers in some states, such as California, Kentucky, and New York, worked to develop alternative assessment models alongside the rise of the high stakes standards, assessment, and accountability regime (Cohen & Hill, 2001; Stecher, 2010). These systems, built around performance-based assessments, were often met with mixed reviews. Some stakeholders emphasized the assessment's ability to investigate student skills that are difficult to quantify, while others criticized the time-intensive demands of such forms of assessment and issues of internal and external validity.

The California Learning Assessment System (CLAS) utilized performance assessments as a measure of accountability for individual students and schools. For English assessments, students answered open-ended questions about texts that they read. For math, students were asked to solve problems and justify their responses. For school accountability, student scores were averaged (Cohen & Hill, 2001).

Political battles forced a quick timeline for the roll-out of the system, and CLAS faced harsh criticism, including questionable validity of the measures incorporated (Cohen & Hill, 2001). Survey results showed that teachers, however, viewed CLAS favorably, which led researchers to suggest that had the system been developed along a more reasonable timeline, it may have had more favorable outcomes.

CLAS provides an example of the importance of the policy climate in which an accountability and assessment system is enacted. The demise of CLAS was influenced by political and media actors' loud criticisms of the policy, even when the attacks were not congruent with what was happening in the classroom (Cohen & Hill, 2001).

The Vermont Portfolio Assessment Program began in the late-1980s and was transformed into the New Standards Reference exam in the late-1990s (Stecher, 2010). Although the later assessments did not require schools to collect portfolios, many districts in Vermont elected to keep doing so. Kentucky implemented the Kentucky Instructional Results Information System (KIRIS), also a portfolio-based system, which faced concerns over the reliability at its start. Portfolio usage and practice varied widely and therefore did not allow for easy comparisons across schools or even classrooms (Koretz et al., 1996). However, like Vermont, Kentucky chose to continue using portfolios at the school level after the end of KIRIS, and an audit in 2008 found they had attained inter-rater reliability of 90% (Stecher, 2010).

A more current example of success with decentralized alternative assessment is seen in the New York Performance Standards Consortium, which was implemented in the 1990s (Knecht, 2007). Not developed via state-wide policy, this approach was developed through grassroots effort. The Consortium was granted waivers for all but one Regent, despite the State's Education Commissioner's preference for standards-based systems (Knecht, 2007). These selected schools now favor a portfolio of performance assessments that students present and defend throughout their schooling over the traditional standardized assessment system. The Consortium is an especially aspirational model because they have determined a way to validate their performance assessments successfully through the creation of standardized rubrics and the use of external reviewers from a board of experts, something they found necessary as they faced

external concerns around validity (Knecht, 2007). This approach holds much promise for those states wishing to replace traditional high stakes assessment systems.

More Recent National Shifts in Assessment Policies

In 2015, Congress passed the Every Student Succeeds Act (ESSA, 2015), which replaced NCLB. With ESSA, the accountability and reporting expectations around standardized testing were somewhat relaxed for states, and alternative assessments in high school were allowed in the policy. Annual testing was still required in grades 3-8, and one test was still required in English and math in high school. However, the stringent accountability sanctions of the NCLB era were removed by ESSA and some states were allowed waivers to "experiment" with assessments other than standardized achievement tests, including performance-based assessments (Evans, 2019, p.3; Rothman & Marion, 2016). This reflected a significant crack in the edifice of federal high stakes accountability and assessment policy prevalent since the 1990s.

More recently and following the COVID-related school closings and learning disruptions of the 2020 and 2021 school years, some state policymakers have supported a reconsideration of approaches and a renewed interest in assessment models centered on student-centered learning. Other state policymakers have retrenched back to traditional standardized approaches of assessment and accountability. Some states, however, have even taken fuller advantage of the flexibility in federal and are finding success with partial or full reliance on performance assessments. At least 14 states, including Virginia, are currently using performance assessments as part of their accountability systems at the classroom, school, state, or federal levels (Alfuth, 2021). Many states, like Virginia, have been hesitant to tie their state accountability to performance-based forms of assessment (Rothman & Marion, 2016). This is likely due to the continued perception of machine scorable exams as the most valid, unbiased, and efficient form

of assessment (Wilson, 2004) even though these assumptions are highly contested (Berliner & Biddle, 1997; Hollingworth, 2009; Nichols, 2007; Randall, 2021; Smith & Fey, 2000). While a majority of states applied for waivers for flexibility in assessment for federal accountability in the 2020-2021 school year, most cited the COVID-19 pandemic as rationale rather than a desire to move to a student-centered alternative form of accountability (McCann, 2021).

Reimagined Systems

In the recent era, scholars of assessment have considered what best practices of assessment and accountability in the 21st century might look like (e.g., Darling-Hammond et al., 2014). One such conception considers an ideal system which would be made of three components: meaningful learning, resource accountability, and professional capacity, and a focus on continuous student improvement. In this version of accountability, the levels of the system that are more removed from students (state, federal) are held accountable for contributing the necessary resources and conditions to schools that are falling short of benchmarks, so that professional educators can do the necessary work to help their students achieve learning benchmarks. This is the opposite of the typical prescriptive or punitive approach that is taken when accountability is focused at the school/teacher level. In terms of assessment, Darling-Hammond and colleagues (2014) advocate for the use of many formative assessments that can inform learning over the use of one summative high stakes assessment, so that more a more diverse range of skills can be assessed, and growth can be emphasized.

There is not a consensus among policymakers and researchers about the use of performance-based assessments for accountability on a state-wide scale. While teachers cited the performance-based assessments as having a more positive impact on classroom instruction, other stakeholders expressed concern over the variability of the assessments, often concluding that

they are unfit for large-scale accountability purposes (Koretz et al., 1996; Koretz, 1998). Others contend that the positive impacts on student learning and increased validity over time outweigh the flaws of portfolio approaches (Stecher, 2010). This is evidence that differing values, such as efficiency, competitiveness, and constructivist approaches, influence policy perspectives. Thus, a discussion about both *what* and *whose* values is necessary as the debates around accountability and testing practices continue. Teachers' preferred values on the topic of assessment and accountability are the focus of paper three.

<div align="center">**Virginia Standards, Assessment and Accountability History**</div>

Definitions

There are a few Virginia-specific policy terms that are helpful to understand when navigating and interpreting standards, assessment, and accountability policy in Virginia. The Standards of Learning (SOLs) are the curricular standards provided to schools that explain the content that must be taught in every subject. These are created by the Virginia Department of Education (VDOE) and approved by the Board of Education (BOE) every seven years, on a rotating basis by subject (e.g., reading, math, history and social studies, etc.). The Standards of Quality (SOQs) are a component of the Virginia Legislative Code. The SOQs mandate and connect these standards to assessments and direct the BOE to operationalize both (Standards of Quality, 2022). These tests are usually referred to as SOL tests. The BOE is also responsible for creating the Standards of Accreditation (SOA) which detail the accountability model used in the state of Virginia. This includes both student- and school-level accountability. For student-level accountability, SOL test scores may be used as part of the decision about grade retention. Additionally, in order to graduate, students must have a prescribed number of verified credits, which are courses in which the student has both passed the course and the SOL test. As of the

2017 SOA, at the school-level, SOL test scores factor into major indicators for the accountability of schools but are not the sole factors.

Introducing the Standards of Learning

The development of Virginia standards, assessment, and accountability policies have largely mirrored the national trends as reflected in the chronology presented in Table 3 below (Duke & Reck, 2003; Heinecke et al., 2003). In Virginia today, the primary accountability system is built around high-stakes tests that assess the SOLs. While the SOLs as they are known today were first implemented in 1995, the initial "Standard of Learning" was introduced at the state level in 1980 (Duke & Reck, 2003). A return to standards in the early 1990s was fueled by political and media debate around the "vaguely-defined values" that were being prioritized by state-level policymakers "at the expense of academics" (Duke & Reck, 2003, p. 42). In 1994, Clinton's *Goals 2000* was implemented, and while the Republican Governor at first thwarted efforts for Virginia to participate in the program, he still supported "higher learning standards" being implemented (Duke & Reck, 2003, p.43).

By the mid-1990s, a new set of standards for Virginia schools had been developed, and state policymakers turned their attention to the ways to measure those standards (Duke & Reck, 2003). Nationally, Virginia was one of the first states to move toward an assessment system based around standardized testing, with the SOL assessments being developed in 1995, before the national requirements of NCLB in 2001.

Table 3

Chronology of Federal and Virginia State Standards, Assessment, and Accountability by Decade

from 1980s-2020s

Years	Federal Policies	Virginia Policies
1980s	U.S. Secretary of Education created NCEE (1981) to address "the widespread public perception that something is seriously remiss in our educational system" (NCEE, 1983, p. 6). *A Nation at Risk* published in response (1983).	First *Standard of Learning* established with long term goal in mind of being able to assess the state *Standards of Quality* (1980).
1990s	Clinton initiative, *Goals 2000,* pushes for standards by incentivizing standards-based reform and testing in states with grant money (1994).	Standards of Learning introduced (1995) and tests piloted (1997) under new Standards of Accreditation (SOA; 1997).
2000s	No Child Left Behind Act mandates high-stakes testing tied to high stakes accountability (2002).	High-stakes SOA start to go into effect that tie student graduation and school accreditation to SOL test performance (2004-2007).
2010s	Every Student Succeed Act (ESSA) requires fewer tests and provides more flexibility but still requires testing in grades 3 through 8 (2015).	Legislative mandate reduces number of SOLs and local performance tasks are introduced (2014).
	ESSA allows for states to apply for waivers from federal testing requirements (2015).	Profile of a Virginia Graduate introduced with focus on 5 C's: citizenship, collaboration, communication, critical thinking, and creative thinking skills (2016).
2020s	Pandemic requires flexibility in ESSA federal accountability (2020 - 2022).	VA allows for flexibility in accreditation during pandemic (2020-2022). VA legislature mandates the consideration of new frameworks for assessment and accountability (2022).

When the first SOL tests were pilot tested in 1997, the scores appeared to show low

achievement, with 2% of schools meeting requirements in all core subjects (Ruff, 2019). While

the BOE expected it to take time for schools to adjust to the new system, the governor of

Virginia quickly politicized the results in a political effort to amplify the need for school

improvement, as a part of his political agenda.

After initial implementation, SOL assessments were tied to an accountability system in

1997, and sanctions were implemented if target goals were not met as specified in the Standards

of Accreditation (SOA) in alignment with requirements of NCLB accountability policy (Duke &

Reck, 2003). At the level of student accountability, the number of credits needed to earn a high

school diploma was increased, and several of those credits were required to be verified with a

passing score on the SOL subject test. Though the system was passed and developed in the late

1990s, the verified credits policy was implemented for the class of 2004 in an attempt to raise the

stakes of the accountability system. Next, individual school accountability was phased in with

the requirement that all schools meet pass rates of 70% in math and reading by the 2006-2007

school year in order to be designated *fully accredited*. For science and social studies, secondary

schools were required to have a 70% pass rate and elementary schools were required to maintain

a 50% pass rate. This accountability system also required schools to share a *Report Card* for

school performance to parents, providing a level of public shame for those not meeting

accreditation standards. Additionally, schools not meeting accreditation thresholds were likely to

have more direct intervention and control from the state. While there was some symbolic level of

input from teachers and district administrators, the legislative policymakers were the primary

architects of the assessment and accountability system (Ruff, 2019). It was a highly politically

driven process.

Profile of a Virginia Graduate

While the SOLs and Standards of Accountability (SOA) dominated the assessment and

accountability policy through the early 2000s, in 2015 with the implementation of the Every

Student Succeeds Act (ESSA, 2015), Virginia policymakers began to decrease the number of

subject areas tested with the traditional SOL standardized tests, especially in science and social

studies. This was part of a shift under which new SOAs were adopted in 2017 and is discussed in more detail in the section on accountability below.

Shift in Standards. In 2014, new legislation passed in Virginia to amend the Code of Virginia and add § 22.1-253.13:3, which would mandate that by 2017 public schools in Virginia would reduce the number of SOLs taken and replace them with authentic performance assessments across science, social studies, and writing, especially in the elementary grades.

The Profile of a Virginia Graduat*e* guidelines (Virginia Department of Education, 2022), developed in 2016 following state legislative mandates described above, indicated that Virginia policymakers were beginning to think about soft skills (i.e., collaboration, citizenship, communication, and critical and creative thinking), in addition to hard content skills assessed by the SOLs. Former BOE member Atkinson (2017) wrote:

> …over the past few years, the Virginia board came to understand that many of its schools had become so laser focused on the state assessment of student performance against the content standards—known in Virginia as the Standards of Learning (SOL)—that they had failed to ensure that students could apply the content knowledge and had developed the skills needed to succeed in the workforce. (p. 28)

In order to address these standards and skills, the introduction of state mandated performance assessments entered the policy discourse, as discussed in more detail in the following CPA.

Performance Assessments. In 2017, following the development of the Profile of a Virginia Graduate and updates to the Standards of Accreditation, Virginia policymakers began to encourage districts and regions to create performance assessments, an assessment type aligned with the goals of the Profile. The rubric scored results of these assessments in writing, science, and social studies were reported to the VDOE; however, from the mid-2010s to the early 2020s,

measures of accountability have remained strictly aligned with the content-based SOL standardized tests. This addition of the Profile policy goals and shifts in alternative assessment procedures were merely laid onto existing accountability and assessment policies and this had led to ambiguity for educators in Virginia school divisions engaged in the accountability system (see paper two).

Concerns About Assessment in the Virginia Context

The concerns about race, class, and misalignment about standardized systems in federal policy previously discussed are equally applicable to the Virginia context. Brunn-Bevel and Byrd (2015) traced how attempts at racial integration and Massive Resistance in Virginia contributed to the disparate outcomes for students in the 21st century: "We find that the locale of schools and their close links to white financial advantage and black student segregation can impact school resources and influence black students' performance on standardized tests" (pp. 419-420). Additionally, as predicted by Campbell's Law, the high stakes accountability system has been marred by cheating (e.g., Evans & Remmers, 2017; Mattingly, 2018).

Recent Political and Policy Changes

There are several recent changes that have been made in the leadership at the state level in Virginia, which have impacted policy shifts in standards, assessment, and accountability over the 2022-2023 academic year. The transition from a Democratic governor to a Republican governor in 2022 came with new layers of political spectacle related to education, which was a major part of his platform while running for office. This can be clearly seen in the rhetoric around his campaign. On March 9, 2023, the governor appeared on a moderated CNN town hall which was entitled "The War Over Education." This choice of semantics shows the elevated political dimension brought to educational policy discussions. At the center of this "war" have

been decisions made about standards, assessment, and accountability. This "battle" is playing out

in a politically split BOE, where half of the appointees were appointed by the previous governor

and half were appointed by the current governor (See Table 3). Some Board members are

appointed to back-to-back terms, the term lengths of the currently appointed members are

included in Table 4.

Table 4

Composition of the Virginia Board of Education July 2022 - March 2023

BoE Members Appointed by Previous Democratic Governor	BoE Members Appointed by Current Republican Governor
President of the Board (2015- 2023)	Superintendent of Public Instruction, non-voting serves as secretary (July 2022- Resigned March 2023)
Vice President of the Board (2019-2023) 2 additional board member (2019-2023; 2017-2025)	2 members appointed to unexpired terms (July 2022-2024; 2022-2025) 2 members appointed to full terms (July 2022-2026) 1 member appointed and removed by the state Senate (July 2022-2023).
As of March 2023: **4 voting members**	July 2022 - February 2023: **5 voting members** As of March 2023: **4 voting members**

Broadly speaking, recent shifts in standards, assessment, and accountability policy in

Virginia have included revisions to and additions of new standards, the introduction of growth

assessments and performance assessments for verified credit, and a change in the accountability

system focused on both proficiency and growth. Much of this has been shrouded in a context of

political spectacle, as discussed in detail in the following CPA section of the paper. Additionally,

in the 2022 legislative session, a law was enacted from House Bill 585 (HB585, 2022) that requires that the Secretary of Education and Superintendent of Public Instruction form a work group to consider new pathways forward for the assessment and accountability system in the state with an implementation timeline through 2028. As accepted, the policy states:

> a work group to revise the Standards of Learning summative assessments of proficiency and to develop a plan for implementation of such revised assessments that shall consider best practices and innovations in summative assessments of proficiency, alternative approaches to current and new assessment items, assessment items that include open-ended questions, long-form writing, and other tasks, a plan for pilot implementation of such assessment items prior to the 2027–2028 school year, the development of a bank of vetted sample assessment items, recommended legislative and regulatory changes and funding necessary to implement approaches considered by the work group, and a proposed timeline for implementation. (HB585, 2022)

How much influence this work group will have on assessment policy changes is yet to be determined. In early 2023, the VDOE staff proposed a new model for accountability that will be considered by the BOE throughout 2023. The intent is for the VDOE staff and BOE to collaborate with the HB585 Work Group in this work (Observation, BOE Meeting, February 1, 2023), but the group is only convening to meet for the first time in March 2023. A member of the BOE is also part of the work group, which should allow for information to be shared across the groups as work is carried out in the upcoming sessions and years.

At the beginning of March 2023, the current Superintendent of Public Instruction resigned, agreeing to continue to serve as a consultant to the Governor's office. This leaves a gap in leadership at this time of transition in the accountability and assessment policy. As new

leadership is appointed and decisions about shifts in the goals and intentions of assessment and accountability models are being made, additional divergent policy intentions may contribute additional confusion about the assessment policies for educators on the ground. This is why a CPA of assessment policy is necessary at this time. With this history in mind, we turn to critical analysis of the system of standards, assessment, and accountability policies within the Virginia policy context.

Standards, Assessment, and Accountability in Virginia: A Critical Policy Analysis

CPA Methodology

The following section of the paper presents a critical policy analysis of the standards, assessment, and accountability system in Virginia from 2014 to the current moment. While these components are interrelated, each component of the system will be considered individually below, with a specific focus on divergence in messaging around the Standards of Learning and the Profile of a Virginia Graduate. Considering each component of the system, overall this study aims to investigate the ways in which differences in policy-in-intention and policy-in-implementation have been constructed, are apparent, and how those differences create ambiguous mixed messages for transitions to policy-in-action and policy-in-experience by educators at the district, school, and classroom levels. The Critical policy analysis presented in this paper focuses on the creation of ambiguity at the state level that becomes problematic for educators' policy interpretations and actions in schools and classrooms. One of the sources of policy ambiguity are the activity of the three bodies of policymakers that inform and define policy: the legislature, the Virginia Board of Education (BOE) and the Virginia Department of Education (VDOE). Table 5 describes each of these policy actors and helps to clarify, "what their specific policy roles and responsibilities are within the educational policy system, how they

interact with other policy actors and networks, and to what end" (Young & Reynolds, 2017,

p.32).

Table 5

Public Education Governance of Virginia K-12 Schools

Policy Actor	Elected Government (Legislature + Governor)	Appointed Board of Education (BOE)	Department of Education (VDOE) including appointed Superintendent of Public Instruction
Responsibilities within the Assessment and Accountability System	Can amend the law that defines responsibilities of BOE & VDOE through the Code of Virginia	Supervision of public school system (§VAC20-131-5)	Evaluation of performance of public schools (§VAC20-131-5)
	Can mandate parameters around the BOE supervision and VDOE evaluation	Decisions by the BOE are often carried out and implemented by staff of the VDOE	Division Superintendents report mandated requirements to VDOE
	Governor appoints members of the BOE and Superintendent of Public Instruction within the VDOE		

According to Guba (1984), there are varying definitions of *policy*, and those definitions

are critical to understanding the policy-in-intention level of policy formation. Table 6 helps

delineate within which definition of policy each of the state-level policymakers are acting. As is

clear in Table 5 and Table 6, while all policy determiners are at the state-level, each policy actor

is making policy mandates for interpretations by policy actors in the next tier(s).

Table 6

Policy-in-Intention: Definitions of Policy and Responsibilities of State Policy Actors

Policy Type	Policy Actor	Definitions of Policy	Proximity to Policy-in-Experience	Policy Looks Like
policy-in-intention	**State Legislature**	–goals of the system –guides –mandates	distant	– prescribed goals for schools and students, parameters for education policy –guidelines for schools – creates rules & work for the BOE and VDOE
policy-in-intention		–goals for students –standing decisions –problem solving	distant	–prescribed outcome goals for students –creates work for the VDOE, divisions, and schools
policy-in-action	**Board of Education (BOE)**		intermediate	–decisions & tactics about how to meet goals, creation of education policy within defined parameters – reviews accreditation status assigned by VDOE
policy-in-intention		–guides	distant	–guidelines for schools and divisions –creates work for divisions, schools, and teachers
policy-in-action	**Department of Education (VDOE)**	–sanctioned behaviors	intermediate	–expectations and evaluations through enforcing BOE rules –grants accreditation through accountability system

Note. Adapted from Figure 1 (Guba, 1984, p.65)

Both political tensions and differences in ideologies of actors between the levels and changes in composition of policy actors within each level over time often impact the interpretation of the policy decisions and the implementation of the policies, as will be seen in the CPA below. The interactions between these policy actors are complex, but broadly, we can define the state legislature, with approval or veto power by the governor, as primary player in creating and mandating goals for the system of education. Legislation they pass creates the

Virginia Code which sets the goals for the BOE and VDOE to interpret. The BOE then is the primary supervisor, interpreting those legislative goals and making decisions about how to meet them. This is done in tandem with the VDOE; however, decisions made by the BOE often provide work for the VDOE as they translate the decisions of the Board into guidelines and expectations in order to evaluate the public school system. This is not a strictly hierarchical relationship, as VDOE staff input is often used to inform BOE decisions and both are influenced by gubernatorial appointments. Of these three policy actors, the VDOE is closest to school-level implementation, meaning that while usually defining policy-in-intention, it also sometimes is enacting policy, especially within the role of evaluating data reported by school divisions.

In considering the roles of these policy actors, this CPA will consider issues of power and how the assessment and accountability policy system itself reproduces social inequities as the policy is experienced. In response to these power and social struggles, examples of resistance by those experiencing the policy will be shared.

Data Sources

This CPA of Virginia's assessment policy was based on a variety of data sources including: policy document analysis, observations of Virginia BOE meetings, and a media review, as summarized in Table 7. Policy documents that were collected included: VDOE policies on the website, the Virginia Code, superintendent's memos and guidelines documents related to standards, assessment, and/or accountability, and publications and executive orders from the Governor's office, totaling to 22 pertinent state policy documents. A review of agendas of public Virginia BOE business meetings and work sessions from January 2019 - March 2023 around relevant policy changes led to 31.75 hours of observations of Board meetings. Additionally, 17 communications videos created by the VDOE and the BOE about the creation

of the Profile of a Virginia Graduate and the accompanying accreditation updates were watched

totaling an additional 1.25 hours of data related to policy-in-intention. In order to gauge history

of and political attention to the policies as well as notions of resistance, a media review of the

Washington Post and *Richmond-Times Dispatch* articles related to standards, assessment, and

accountability from January 2016 to early March 2023 were collected, totaling 1,658 articles.

These headlines were read and narrowed to the 224 articles directly relevant to the topic of this

study. One last media data source was used, the CNN Town Hall with the governor of Virginia

titled the "The War Over Education," which aired in March 2023.

Analysis Procedures

As advocated for by Young (1999) when conducting CPA, memoing, critically reflective

journaling (Janesick, 2016), and the use of critical partner dialogue were used for all of the

analysis procedures. For the policy documents reviewed, the data were analyzed by thematically

coding the documents using each of the tenets of CPA as a priori established codes (see table 7

below). Additionally, the policy documents were coded for each reference to the Standards of

Learning and SOL tests and the Profile of a Graduate and associated performance assessments.

This coded data was then used within the analytic induction (Erickson, 1986), described in detail

below.

For the media review and observations of recorded Board meetings, ethnographic content

analysis (ECA) procedures (Altheide & Schneider, 2013) were applied. This was a relevant

method because "ECA is strongly oriented to qualitative data analysis, which involves

description, attention to nuances, and openness to emerging insights. ECA acknowledges the role

of reflexive observation as one becomes immersed in relevant documents, but the aim remains to

identify similarities and differences in how the documents—or parts thereof—reflect other

aspects of …the cultural context" (Altheide & Schneider, 2013, p. 27). ECA combines

theoretical and saturation sampling in order to capture the data relevant to the topic of interest.

This was applied to both Board meeting observations and the media review.

Analysis of Board meetings required purposive sampling by looking through all agendas

for sections relevant to standards, accountability, and assessments followed by finding the

corresponding components of the cataloged video for BOE meetings and work sessions. The

selected and observed segments totaled 31.75 hours of observations. Field notes were written

during observation. Additionally, analytic memos were written iteratively throughout the

observation process.

For the media review, all headlines were placed into categories that were iteratively

developed during data collection and are reported in Table 7. While all news articles were

skimmed, studies were purposively sampled and included in their totality in the data based on

relevance to the assertions generated, as recommended by Altheide & Schneider (2013).

In order to generate assertions from the data corpus, framed around the analytical

concepts of CPA, methods of analytic induction (Erickson, 1986) were employed. This method

of analysis requires a repeated reading of the data corpus, through which analytic assertions were

developed. Assertions are claims that the researcher constructs based on the data as a whole,

which are warranted through a reading and rereading of the data corpus and a search for

disconfirming evidence. This process was conducted to assure that the data support the

assertions. Three main assertions were developed: 1) There are differences between policy

rhetoric and practice, often spurred by political and ideological differences between

policymakers; 2) This creates mixed messaging for those interpreting and implementing the

policies; and 3) Mixed messages slow intended progress, even when there is advocacy for change.

Table 7

Data Sources and Analysis Procedures for CPA

Research Questions Explored	1. What is the rhetoric and values surrounding accountability and assessment policy in the state of Virginia? a. Where does mixed policy messaging exist? b. How might the conflicting values and intentions embedded in current assessment policy create ambiguity for educational practitioners? 2. Where are notions of resistance occurring that indicate divergences in policy from practice to subvert oppressive systems?		
Data Source	**22 Policy Documents** – VDOE policies on the website –the Virginia Code –Superintendent's Memos –Guideline Documents –Governor's related executive orders and publications	**33 Hours of Observation** – BOE Work Sessions (2019-2023) – BOE Business Meetings, including public comment (2019-2023) –17 Official VDOE/BOE communications via YouTube (2016-2018)	**224 Articles via Media Review (2016-2023)** –Including letters to the editor and quotes from resisting citizens
Source Specific Analysis Procedure	Thematic analysis using codes: -Accountability –Assessment –Mixed Messaging –Notions of Power –Policy Origin or History –Profile of a Graduate/5 C's –Resistance –Rhetoric vs Outcome –Social Stratification –SOLs –SOL tests	Ethnographic Content Analysis (Altheide & Schneider, 2013) using categories: –2016/2017 reform –accountability and accreditation –assessment outcomes –history standards –innovative ideas –pandemic assessment decisions –political coverage related to education –SOL cheating –concerns about race related to education –resistance of the current system	
Overall Analysis Procedure	memoing, critically reflective journaling (Janesick, 2016), use of critical partner dialogue, analytic induction (Erickson, 1986)		

Validity of the study was enhanced through triangulation from multiple data types and sources, as analytic induction asserts that data from a single-source is less robust. In the reading of the data corpus, assertions were edited and are presented below with evidence and interpretive discussion. Results of the CPA presented below are organized by category: standards, assessment, and accountability, and assertions will be addressed throughout for reader ease. The following results summary section will summarize how these results match the assertions.

<div align="center">**Results**</div>

Virginia Standards

Legislative Policy

As discussed in the opening history, the Standards of Learning (SOL) and standards-based reform were first incentivized via political and governmental forces in the mid-1990s, and ties of standardized testing to accountability were first mandated through the federal NCLB in 2001. The policy shift to focus on thinking about necessary 21st century skills needed by students was also initiated by legislative forces, this time at the state level in the mid-2010s.

In May 2016, the governor signed into law a mandate for the Virginia Department of Education (VDOE) to create the Profile of a Virginia Graduate (Llovio, 2016). The legislation was, in part, created by the *SOL Innovation Committee,* which is a partnership between non-partisan legislators and educators. The enacted legislation mandated:

> the Board shall develop and implement, in consultation with stakeholders representing
> elementary and secondary education, higher education, and business and industry in the
> Commonwealth and including parents, policymakers, and community leaders in the
> Commonwealth, a Profile of a Virginia Graduate that identifies the knowledge and skills
> that students should attain during high school in order to be successful contributors to the

economy of the Commonwealth, giving due consideration to critical thinking, creative

thinking, collaboration, communication, and citizenship. (Standards of Quality, 2018)

Within the Virginia Constitution, the power is given to the Virginia Board of Education (BOE) to

"promote the realization of the prescribed standards of quality" in a way that is "subject to such

criteria and conditions as the General Assembly may prescribe" (art. VIII, § 5). Of note here is

the focus on the need for students to be *successful contributors to the economy*, indicating that

when legislators are thinking of the purpose of public schooling, they prioritize teaching skills

necessary for economic mobility, or for students to be informed consumers, as a primary goal.

Policy Interpretation: Virginia BOE and VDOE

The BOE writes and adopts the required "Standards of Quality," often informed by

legislative mandates, and the Standards of Accreditation (SOA), which are the official written

policy of the accountability system for Virginia public schools. This policy is written in the

Virginia Administrative Code (hereafter the Code), Title 8 which pertains to education, under

agency 20, the State Board of Education, Chapter 131. With the legislative mandate given in

2016, new SOA were adopted in November 2017 to be implemented in the 2018-2019 school

year.

Standards, assessment, and accountability are interwoven in their presentation throughout

the SOA, however mentions throughout help indicate intended policy preferences and values. As

of March 2023, throughout the entire section of the Code relating to public schools, the

Standards of Learning and SOL tests are mentioned more times than the Profile of a Virginia

Graduate and performance assessments (See Table 8 below).

Interestingly, in the sections about *purpose* skills related to the Profile of a Graduate are mentioned, but references to content measured by the Standards of Learning are scant. For example:

> The foremost purpose of public education in Virginia is to provide children with a high quality education giving them opportunities to meet their fullest potential in life. The standards for the accreditation of public schools in Virginia are designed to ensure that an effective educational program is established and maintained in Virginia's public schools. The mission of the public education system is to educate students in the essential knowledge and skills in order that they may be equipped for citizenship, work, and an informed and successful life. (Standards of Accreditation, 2018)

While the section on purpose does mention that schools should be accredited, which currently is largely tied to the SOL scores, the only skills mentioned are those relevant to the Profile, such as citizenship. As discussed in depth below, the skills for work mentioned in the SOA are considered by the BOE as more closely aligned with the 5 C's than the SOLs. With accountability being tied to one outcome measure, SOL tests, and purpose being tied largely to another, the *Profile*, this is an example of a mixed messaging by the state policy actors. Additionally, here the Board takes a different interpretation of the purpose of schooling than the legislature. While the legislature focused on economic outcomes, the BOE rhetoric indicates that the Board is interested in preparing citizens for democratic society, in addition to preparing people for the workforce and for a more holistic conception of a successful life.

Table 8

Standards of Accreditation (2018) Mentions of Policy Mandates

Policy	Number of Mentions
Standards of Learning	21
SOL tests	26
Profile of a Virginia Graduate	7
Performance Assessments	6
5 C's mentioned together	7
One or more skills associated with 5 C's mentioned individually	11

More mixed messaging is included in the section of the Code about philosophy, goals, and objectives of public schools. Skills mentioned in the 5 C's are highlighted, "their education should encompass mastery of creative and critical thinking, analysis and problem solving, and the development of personal attributes such as communication and collaboration skills, dependability, and persistence," before any mention of the Standards of Learning (Standards of Accreditation, 2018). This is evidence that there is misalignment and mixed messaging between the priorities of the public school system. While the purpose and philosophical goals are more aligned with skills and process, the measured outcomes and evaluation are more aligned with content and standardized assessments. We will see how these changes have transpired in the following section.

Board Priorities

The BOE clearly stated their priorities for 2018-2023. On their page within the VDOE's website they identify three areas which are the most important to them:

Priority 1: Provide high-quality, effective learning environments for all students

Priority 2: Advance policies that increase the number of candidates entering the teaching profession and encourage and support the recruitment, development, and retention of well-prepared and skilled teachers and school leaders

Priority 3: Ensure successful implementation of the Profile of a Virginia Graduate and the accountability system for school quality as embodied in the revisions to the *Standards of Accreditation* (VDOE, 2023b)

These priorities were adopted as part of the Virginia BOE Comprehensive Plan that covers the years 2018 to 2023 by members of the Board in conjunction with the Superintendent of Public Instruction at the time. While the priorities are assigned through 2023, other messaging from the Board and Department seem to indicate that priorities have not been consistent over time. The priorities were very much aligned with the priorities of the Board at the time of implementation.

VDOE and BOE Policy Activity in 2017-2018

The first interpretation of the Profile of a Virginia Graduate policy-as-intended is rendered through promotional videos produced by the VDOE when the Profile was introduced. Through a series of five videos, the DOE introduced the new SOA, including the Profile and the 5 C's. In the introduction video, the Superintendent of Public Instruction at the time noted that the previous accountability system helped to build continuity across the state, gathered school data, and helped schools to make data informed decisions. He also recognized that there were gaps when the sole focus is on the outcomes of the SOL tests. He noted that SOL tests will

remain important but that "they're not completely adequate" and that this new approach would

allow for a broadening to a more holistic approach to accountability (VDOE, 2017c). He asserted

that the policy goal at the school-level for the new SOA was to ensure that schools were

continually improving. At the student-level the goal was to:

> Create opportunities for students to access 21st century skills. Skills that we heard
>
> repeatedly are important for employers, the military, for 2 year, and 4 year education
>
> outcomes. The notion is that knowing things is important. Content knowledge is
>
> important. But we have to go beyond that for students to have successful opportunities in
>
> the 21st century. So we're talking about communication skills, the ability to collaborate,
>
> creative problem solving, critical thinking, and engaged citizenship. (VDOE, 2017c)

This quote helps illuminate that the VDOE sees the SOLs and the *Profile* as compatible, with

shared importance between content knowledge and skill development. He goes on to explain,

that one mechanism of creating opportunities for students to access the 5 C's skills is by:

> reducing the number of verified credits that we use SOL tests for. That affords students
>
> time to explore other types of assessment and other opportunities. So they might use an
>
> assessment by performance standards, where they're making a presentation, where
>
> they're writing a paper, so they can demonstrate those important skills in that way. The
>
> intent of these reforms, both in accountability and around the Profile of a Graduate, are
>
> to make sure that Virginia schools, all of them, are preparing students in a way that
>
> creates opportunities for a high quality life in the 21st century and encourages our
>
> schools to continue to get better. (VDOE, 2017c)

Verified credits are courses where the student is required to pass the course and prove mastery

with the passing of a state standardized test. The SOA sets the number of verified credits that are

required for graduation, and reduction of required verified credits meant that the number of SOL tests required was reduced. This quote indicates that performance assessments are more aligned with demonstrating the skills needed for a "high quality life" (VDOE, 2017c).

In another video in the series, entitled *The Case for Reform*, a VDOE staff member and BOE member noted that when the SOLs were established, Virginia was at the low end on NAEP comparisons. In order to remedy this, the SOA in the 1990s were updated to include common state content standards, the SOLs, that were assessed with SOL tests. They added an accountability model to accredit schools that was punitive in nature. While a few minor changes have been made, this has been the basic standards, assessment, and accountability system since first implementation. In the video, they explain a major weakness of the assessment system was that it did not measure growth:

Our schools have made a great deal of progress but in many instances our system doesn't recognize that progress. Certainly, we have really improved our NAEP scores and there's no doubt about that, but when we look at how we were looking at schools, we really were looking at an aggregate group of students, and if they met the benchmark…then a school was accredited, but we sort of left out students who started well below grade level and made over a year's worth of progress but still did not pass the test. We were not giving schools credit for that, and I think that is one of the unintended consequences of our program. We were not able to recognize that progress, and I think it doesn't give the public a true picture of what is really occurring in schools. (VDOE, 2017a)

This rhetoric here is important. The main weakness was that when looking at total proficiency, students who were making great strides of progress but still remained below the benchmark were invisible. As will be discussed in upcoming sections, this idea of being invisible within the

accountability model is a motif that will resurface. The VDOE and BOE identified a few additional unintended consequences of the traditional system that are mentioned in the policy discourse. For instance, a Board member noted that the development of the SOL system led to significant emphasis on test preparation and a narrowing of the curriculum:

> I don't think folks ever anticipated when the program was put in place the emphasis that would be put on preparing for the test. And … the additional skills that teachers might focus on, but they're not, because they're really focusing on content alone, and so the absence of skills focus is another area that we have concerns about. (VDOE, 2017a)

BOE members and VDOE staff in this era indicate that 5 C's skills should be a greater focus in the classroom; however, due to the remaining standardized tests, there is evidence that the curriculum was narrowed more significantly and misinterpreted by policy actors at the school level (See paper two).

DOE and BOE Policy Activity in 2022-2023

While there was a movement afoot by policymakers to be less tied to the Standards of Learning in the mid-2010s, VDOE staff in 2022-2023 have indicated solid support for the SOLs. This is reflected on the VDOE website in 2023. As seen in Table 9, while there are 116 pages of the website that reference the Standards of Learning, only 10 pages include mentions of the Profile of a Virginia Graduate. While one may think this could be due to the long policy history of support for the traditional SOL assessment system and the only more recent introduction of the Profile, this does not include archival documents stored on the VDOE website, only active pages with information for parents, students, teachers, and administrators. This is evidence that at the VDOE, the SOLs are prioritized.

Table 9

VDOE Active Pages with Mentions of Policy Mandates in March 2023

Policy	Number of Mentions
Standards of Learning	116
Profile of a Virginia Graduate	10

New Standards in History and Social Studies

The mixed messaging and tension from the competing SOL standards and 5 C's skills have been especially evident in the legislatively mandated routine effort for the BOE to approve the updated Standards of Learning for History and Social Studies (H&SS) in 2022 and 2023. The controversy over what should be in the standards illustrates that standards policy is a highly politicized process. Some have considered the revisions white-washing while others have considered it leftist indoctrination (Ayers, 2023; Bryson, 2022b). Additionally, and very much related to the notions of the SOLs and the 5 C's, paradigmatic pedagogical differences have been apparent in ongoing Board debate about whether the standards should be more centered on facts or skills. As we know from Rein (1976), politicians are apt to develop ambiguous policies in order to compromise in legislative and quasi-legislative processes. The embedding of multiple values and meanings of such policies allow, and even require, for ambiguities to be worked out in lower levels of the policy-in-action and policy-in-experience levels of the system where additional value conflicts can occur in the reinterpretation processes. These ambiguities in standards setting will also manifest in mixed messages in the system of assessment discussed later in the paper.

Virginia law mandates that all of the Standards of Learning are reviewed every seven years. During the previous Governor's tenure (2018-2022), H&SS SOLs came up for review in the regular cycle in 2021. At that time, the VDOE staff presented a timeline and process for revising the H&SS standards which was approved by the BOE in January 2021. That process was carried out and a draft of revised standards was presented to the BOE for first review in August 2022. At that meeting, the VDOE staff person who had led the entire revision process explained that the new development of the H&SS standards had included considerations of the Profile of a Virginia Graduate and the 5 C's. The previous revision of the H&SS standards had occurred prior to the Board's adoption of the Profile. In the presentation, there was clear indication of what their legislative responsibility was:

> The Board of Education shall establish educational objectives known as the Standards of Learning, which shall form the core of Virginia's educational program, and other educational objectives, which together are designed to ensure the development of skills that are necessary for success in school and for preparation for life in the years beyond. (Standards of Quality, 2021)

From this legislative mandate, the VDOE staff person explained the vision of the VDOE staff involved with the two-year revision process, as well as educators who were involved:

> One of the big things that came out of those discussions with those practitioners was that we want to value history and social science as more than just content that we memorize, but content that our students can do something with for post high school success, and I really like what we have here in the sense that the standards are more than just passing a test. We're looking at the standards…to honor our collective history, really taking into

account that no single narrative is going to be enough. We really need to look at those

multiple perspectives." (Observation, BOE Meeting, August 17, 2022)

Embodying the vision from multiple perspectives, the revision process involved input from over

5,000 public comments and over 200 additional contributors, including educators, historians,

students, parents, business leaders, museum employees, and administrators. The synthesis of all

these ideas was applauded as an exemplary process by the entire BOE (Observation, BOE

Meeting, August 17, 2022). Input from the revision process resulted in the decision to make two

pedagogical shifts to the standards. The first was to consider topics more thematically instead of

chronologically. The second was to engage with the themes through inquiry-based practices, a

model aligned with the Profile and the 5 C's. The end product was a 400 page document that

included: each standard, the content that students should understand within that standard,

supporting questions to promote inquiry-based approaches, and recommended learning

experiences to support teachers as they develop lessons for the standards. In the past, the

standards; the curriculum framework, which included content students would be tested on; and

an enhanced scope and sequence, which included support for lesson development, were three

separate documents.

In discussion following the presentation, the newly appointed Superintendent of Public

Instruction, appointed by the new conservative Governor, pushed for a one month delay in

consideration of the new standards. She said that the two main reasons she wanted to not accept

the standards for first review were because of time and needed corrections, citing that most of the

needed corrections were typographical errors. In terms of time, she cited that she wanted the

newly appointed Board members to have more time to review the standards. The president of the

BOE pushed back on this recommendation for delay noting that the VDOE, under her leadership,

had the document for seven months and that the Board was not responsible for reviewing for typos. While all Board members agreed that the document was not ready for final review, the long-standing BOE members wanted to accept the document for first review to allow for minor changes to be worked out within the revision process, while the newer members did not, likely due to their want for sweeping changes. One newly appointed Board member raised his concern about the content saying, "Facts matter. Memorization matters. Repetition is important" (Observation, BOE Meeting, August 17, 2022). The President responded that if the request was only for a month's delay, that "going back to fact-based" was going to take longer than that (Observation, BOE Meeting, August 17, 2022).

Another newly appointed board member stated that he thought the standards and idea for inquiry were good, but that there was also a need for more content. He noted that he believed the media would portray this argument as left/right or woke/not woke but that his concerns about the standards "run the gamut" (Observation, BOE Meeting, August 17, 2022). Another newly appointed Board member noted that "this is not political" but that "history is chronological" and while themes are important, she had concerns about the thematic presentation. After this debate, the BOE agreed to delay the first review.

The following month, a Board member raised concerns about the presentation of the Profile of a Virginia Graduate within the H&SS standards document. She thought it's inclusion and explanation made the document "hard to navigate;" additionally, she raised her concern about this skills-oriented approach to the content:

> One major point that I have is that the schools are advocating for what's called inquiry learning, and it's known by many names: discovery learning, problem based learning. experimental learning, constructivist learning, learning experiences, and the processes, it

is appealing to some, but empirical research has shown that, you know, overwhelming unambiguous evidence that minimal guidance during instruction is significantly less effective and efficient…to introduce something called an inquiry learning at this time…I think it's backward. (Observation, BOE Work Session, September 14, 2022)

This statement illustrates the lack of support by a Board member for the inquiry-based approach well suited to the 5 C's, Profile, and performance assessment. The VDOE staff person presenting asked for the BOE member to share the research that she was referencing, because the majority of research that the staff person was aware of supported the inquiry approach. This illustrates conflict on the BOE and in the VDOE about the standards and assessment approach, which is a source of policy ambiguity. These interactions reflect how standards and assessment policy is political. At this meeting, they asked for the standards to be separated from the rest of the content within the 400 page standards document in order to more easily present just the standards to parents and the wider community for public comment.

In November 2022, a new draft of the standards was brought to the Board by the Superintendent of Public Instruction, which the Board rejected (Observation, BOE Meeting, November 17, 2022). This rejection was characterized as politically motivated because the original draft included input from academics who were considered left-leaning while the November draft's contributors were from members of conservative think-tanks, and this draft was considered right-leaning (Bryson, 2022b; 2022c). In the BOE meeting, there were several reasons why the Board ultimately rejected the changes: a) the document had broad changes, some of which were characterized as white washing, that were different from the separation of the standards the BOE had asked of the VDOE at the previous meeting, b) the new standards were delivered within a very short time for reviews, c) there were many errors, which the

Superintendent of Public Instruction acknowledged and characterized as accidental. Concerning

the controversy over releasing the new H&SS standards, one board member stated in a BOE

meeting:

> I'm not sure where we're going from here because I don't believe that this Board ought to
> legitimately put out this document as the work of the Board given the length of time, or
> lack thereof, we've had to review it, and frankly given the absence of transparency with
> regard to how it got to where it got from where it was, when the direction was separate
> out the framework from the standards. (Observation, BOE Meeting, November 17, 2022)

This shows that standards development is a highly partisan political process that leads to conflict

and ambiguity in policies for local educators to interpret. It sends no clear message of support for

either the SOL or Profile of a Virginia Graduate, or the assessment approaches affiliated with

them.

There was also concern by Board members over components that had been deleted from

the standards, including mentions of the Profile of a Virginia Graduate. One board member

stated, among a list of many missing components within this version of the H&SS standards that

the Profile had "disappeared" (Observation, BOE Meeting, November, 17, 2022). While this

Board member was concerned that the Profile was gone, in the previous meeting other BOE

members had expressed their negative concern about inclusion of the Profile. This shows how

even the level of interpretation at the state level is ambiguous, with the BOE's mixed messaging

from different members complicating the translation to policy process and final documentation

for the VDOE.

A major concern with this November 2022 draft of the standards was that, unlike the

August draft, there were many topics previously included in the last iteration of the standards, in

2015, that were removed from this November draft. A BOE member shared her concerns about these deletions, which she indicated that she considered an attempt to whitewash the curriculum:

> I have defended the administration's handling of this matter, both publicly and privately. I have told people who thought that this was all an attempt to whitewash 'No, no, no! Calm down! We're really just trying to get it right!' I no longer have that confidence. I can no longer say that to those folks….[the VDOE] told us that Martin Luther King and Juneteenth were *accidentally* left off the list. I understand that [the Superintendent] has acknowledged today that the 'first immigrants' reference [in reference to Indigenous people] was also an accident. The nature of the accidents…are extreme….Simply even comparing the November draft to the 2015, the existing standards. It is a major step backwards.…The 2015 existing standard, in our mini world history, have us study Greece, Rome, the empire of Mali, China, and Egypt. It's now Greece and Rome…. Africa and Asia have disappeared from the elementary school curriculum. There is no discussion of Africa or Asia in the early [grades]...in third grade we talk about European geography. It's *so* Euro-centric. It's just mind boggling." (Observation, BOE Meeting, November, 17, 2022)

Her claims of whitewashing in the quote were furthered when she additionally noted concerns with the second grade curriculum, where Martin Luther King, Jackie Robinson, Thurgood Marshall, and Cesar Chavez, who used to be studied as leaders of the nation, were no longer included in the standards. So, while previously, a newly appointed conservative Board member indicated that the reason to delay first revision of the standards was not due to a "woke/not woke" argument (Observation, BOE Meeting, August 17, 2022), there were more changes to the

document in the November 2022 presentation of the standards than simply the editing for clarity that was previously agreed upon by the Board.

 In January 2023, a new draft of the H&SS standards was released in an attempt to bridge the two versions. The Board approved the draft at their February 2023 meeting and solicited input in a series of public comment sessions scheduled around the state in March 2023 that will inform their next draft.

It's important to note the context for this debate, which has received much more media attention than SOL standards reviews usually receive, and more news articles than any other single education topic in the time frame of this study's media review. After the new Governor's election, his first act was the creation of an executive order that banned "inherently divisive" concepts and the teaching of critical race theory in schools, which had been a major part of his campaign platform. In conversations about the H&SS standards, there has been a recognition to teach "all history, good and bad" quoted repeatedly in the press, at Board meetings, and at the governor's CNN town hall, but there has been concern about the conflation of African-American history and critical race theory.

As applicable to the standards, there was much Board discussion around the guiding principles that served as an introduction to the proposed standards, which included statements such as: "Teachers should engage students in age-appropriate ways that do not suggest students are responsible for historical wrongs based on immutable characteristics, such as race or ethnicity" (Virginia BOE, 2023, pp. 8-9). A member of the American Historical Association, who helped with the original draft but was not invited to contribute to the revision, stated, "That's a red herring. Teachers are not doing that. There is no reason to say that other than to score points" (quoted in Mirshahi & DeFusco, 2023). This is an example of how political

spectacle can add to the mixed messaging around education, especially within standards, assessment, and accountability policies.

While some board members were very uncomfortable with the front matter that included this quote being included with the standards, others felt that it was an important component of the document (Observation, BOE Work Session, February 1, 2023). The front matter was included in the distributed draft standards for public comment.

The Politics of Standards Policy in Virginia: Race in the Classroom. Part of the controversy around the updated H&SS standards has certainly centered around race, as seen in the Board members extended quote above where she shares her concern about how the early elementary standards have changed. During the CNN town hall, the Virginia governor noted that during the pandemic when students were at home, that parents realized "what was being taught in the schools was pitting children against one another, based on race, or sex, or religion." (Youngkin, 2023). In making this statement, the governor added more ambiguity to the standards policy discourse. He stated that he wants to teach Virginia students both the "good and bad" history and at the same time claims "the inherently divisive concepts," which he banned in his first executive order as governor, "are taken directly from the Civil Rights Act. And they're teaching children that they're inherently biased, or racist, because of their race, or their sex, or their religion" (Youngkin, 2023).

When considering CPA's tenets about notions of power and social stratification, here is a clear indication that power controls what knowledge is considered appropriate and what is not. This use of the Civil Rights Act to make claims of reverse racism is both a common, and commonly criticized, practice (Lipsitz 2006). The banned discussions in Virginia schools have been about how privileges from the past have led to differences in opportunities. This color-

evasive approach, also sometimes called color-neutral or colorblind ideology, which is supported by the governor, is a tactic that has repeatedly shown to replicate inequities (Bennett, 2019; Goldin & Khasnabis; 2021; Randall, 2021; Wilt et al., 2022).

Resistance

Students, parents, and educators, however, are resisting the governor and his administration's political work to manipulate what is happening in schools. At public hearings around the revision of standards, Virginia students have spoken in favor of learning a more accurate history and parents have overwhelmingly said that they support the professional opinion of teachers (Ayers, 2023; Observation, Public Comment at BOE Meeting, November 17, 2022). Teachers have shown up to public comment sessions in support of the original August 2022 standards (Observation, Public Comment Session, March 15, 2023). Additionally, during the town hall, there was significant push back and fact checking about the governor's claims, as collected and documented by Moore (2023).

While much of the resistance is around the SOLs, there is also citizen support for the Profile. Business owners have advocated for support of the promotion of the 5 C's over such a stringent focus on standards. Referring to the 5 C's as "durable skills," a Virginia CEO writes, "It's unsurprising that these skill sets also are in high demand among employers. For Virginia to continue recovering and thrive in a postpandemic economy, it is imperative that the education and business communities work together to prioritize the development of durable skills in our student body" (Nomberg, 2022). He notes that increasingly companies are hiring based on skills over educational attainment and encourages Virginia policymakers to take this into consideration as policy decisions are made. This helped substantiate legislators' claims during the formulation of the Profile of a Virginia Graduate that a purpose of the policy shift was to help promote

economic mobility for students and the Commonwealth. Business leaders seem to support the Profile and its authentic assessment approach over the SOL content-based standardized assessment approach. This further illustrated how multiple and conflicting values may be embedded in policy.

Additional Mixed Messaging: Competing Assessment Systems

If there is ambiguity and multiple messages in standards policy, there will be ambiguity and multiple messages in assessment policy. As we have seen over recent years, there has been accretion in standards policy, adding the Profile and 5 C's to the traditional SOL standards. The Profile was developed because feedback from post-secondary educators and employers noted that Virginia graduates were missing crucial critical thinking and communication skills. How this has affected assessment policy is the subject of this next section. As assessments are based on standards, it follows from the previous discussion on standards that mixed messaging about systems of assessment are also likely to be ambiguous as policy in intentions transition to enacted policy.

Legislative Intent

The Profile of a Virginia Graduate is tied to graduation within the state Code, as described in the standards section of this paper. Standards of Learning, then, in the law are repeatedly referred to in regards to testing. The General Assembly and the BOE, together, adopt the Standards of Quality (SOQ). There are twelve mentions of Standards of Learning assessments throughout the SOQs, and no mentions of alternative or authentic assessments (Standards of Quality, 2021). However, there is also nothing within the Code that prevents alternative assessments from serving as a Standards of Learning test for accountability.

Policy as Written: The BOE's Standards of Accountability (SOA)

In the BOE's SOA, related to student achievement, SOLs are prioritized as the main measure of student outcomes. It is not until the subsection regarding graduation requirements that performance assessments related to the Profile and 5 C's are even mentioned within the achievement section of the SOA. While the SOQs did not define the Standards of Learning assessments, the SOA do. The tests for the Standards of Learning are referred to solely as "SOL tests" while every mention of performance assessments refers to them as "authentic" assessments. In their own definition sections of the SOA, the BOE policymakers make the following differentiation:

> "Standards of Learning tests" or "SOL tests" means those criterion referenced assessments approved by the board for use in the Virginia Assessment Program that measure attainment of knowledge and skills required by the Standards of Learning….

> "Authentic performance assessment" means a test that complies with guidelines adopted by the board that requires students to perform a task or create a product that is typically scored using a rubric. An authentic performance assessment may be used to confer verified credit in accordance with the provisions of §VAC 20-131-110. (Standards of Accreditation, 2018)

This begs the question, if the policymakers see performance assessments as the only authentic assessment in terms of preference and legitimacy, or if they are simply attempting to delineate between authentic performance assessments and other types of performance assessments.

While the goals and philosophies of public education in the SOA were centered on the skills associated with the Profile, in the section entitled "School Instructional Program" the primary attention goes to curriculum as aligned with the Standards of Learning with article A

stating, "As required by the Standards of Quality, each local school board shall develop and implement a program of instruction for grades kindergarten through 12 that is aligned to the Standards of Learning" (Standards of Accreditation, 2018). Then, in article B, the BOE policymakers note that students should attain the skills "in accordance with the Profile of a Virginia Graduate" including the 5 C's (Standards of Accreditation, 2018). So, in essence, there are mandates for dueling systems of content-standards versus skills that require significantly different instructional and assessment approaches.

These policies are then presented with different emphasis for different school levels. While SOLs are the focus for elementary and middle school, with zero mentions of the Profile of a Graduate or the 5 C's, they become prioritized in the direction for the instructional program for high schools as evidenced in the SOA:

> "The secondary school, in accordance with the Profile of a Virginia Graduate approved by the board, shall provide a program of instruction to ensure that students (i) attain the knowledge, skills, competencies, and experiences necessary to be successful in the evolving global economy whether immediately entering the world of work or pursuing a postsecondary education and (ii) acquire and be able to demonstrate foundational skills in critical thinking and creative thinking, collaboration, communication, and citizenship in accordance with §VAC 20-131-70 and the Profile of a Virginia Graduate." (Standards of Accreditation, 2018)

In this section on high school standards for instructional emphases, the Standards of Learning receive zero mentions. It isn't until a further section concerning "verified credits" that they are mentioned again, when at that point, the SOLs are the focus of the policy with only one mention of "authentic performance assessment" in relation to the writing subject area.

Here, we see clearly the mixed messaging being given by the BOE, especially to those actors at the VDOE, district, and school levels within the education system. While the purpose and philosophy of the schools prioritized the 5 C's, as discussed in the standards section, graduation requirements in the accountability standards are still by and large attached to the standardized SOL assessment requirement.

Negotiating Interpretation: The BOE and VDOE's Take on Assessment

As with the standards, the BOE's assessment policy perspectives have changed over time. In the mid-2010s, there was a shift toward performance assessments and away from standardized testing as a priority, with rhetoric about increased freedom and flexibility. By the early 2020s, shifts in policy, that included Balanced Assessment Plans discussed in detail below, show that policymakers have retrenched to support the standardized SOL tests and are concerned less about flexibility than was originally intended by the policy in the 2010 era.

2017-2018. In the 2017 communication video about the new SOA, the *Case for Reform*, members of the BOE and VDOE discuss how in conversations with business and colleges, they learned about the need for a greater focus on skills. The BOE member shared:

> Our students, while they may have content knowledge, what they are missing are some
> soft skills….one of the negative, or the unintended consequences, of our old system was
> that the testing really drove out the emphasis on those skills, and so we want to create the
> flexibility and the opportunity within the schedule for emphasis on those skills as well.
> (VDOE, 2017a)

When asked how they plan to make that space, the answer was clear: "the main way we're doing it is by eliminating the number of verified credits, or the number of SOL tests the students need

to pass in order to graduate" (VDOE, 2017a). These verified credits, part of the accountability system, will be discussed further below.

In 2014, prior to the development of the Profile of a Graduate, the legislature enacted a law which eliminated five SOL tests in elementary and middle school and required that they be replaced with authentic assessments (Standards of Quality, 2014). Three social studies tests, two science tests, and one writing test fell under the purview of the legislation. One of the sponsors of the bill noted that "Instead of filling in circles on a multiple-choice test, authentic assessments ask students to demonstrate the depth of their understanding through essays or projects. Often, the assessments are set up so that students are performing tasks that reflect practical challenges in the real world." (Chandler, 2014a). The rhetoric of this language "filling in circles" versus "depth" indicates a strong preference for the new approach to assessment and the intention to push assessment in a more authentic direction. Along with this decision to cut-back on the number of SOLs tests, the enacted legislation required school divisions to develop local performance assessments for measuring learning in these subjects.

So, in 2015, the VDOE asked school divisions to create performance assessments in elementary and middle school grades to assess the writing, science, and social studies content no longer assessed by standardized SOL assessments. This was, again, a politically motivated change. The democratic governor at the time had campaigned about assessment reform and after signing the 2014 bill that reduced the number of required SOLs into law, called it a "meaningful step toward reforming the Standards of Learning so that we can continue to evaluate students and teachers without stifling innovation and creativity in the classroom" (Chandler, 2014b). This bill also created the *SOL Innovation Committee*, which, as mentioned previously, helped create the Profile of a Virginia Graduate.

Thus, by 2018, the notion of performance assessments had been tied to the Profile and the 5 C's. In a 2018 video released by the VDOE, the performance assessment coordinator shared with schools:

This is a really large shift after 20 years of multiple-choice testing. What we're really headed for in Virginia is a balanced assessment. You may look at some content and know that the only way to test it is multiple choice. You may look at some content and say 'well, that requires some writing by students.' You may look at other content and say, 'Well that would lend itself to a curriculum-embedded task.' Like in science, we should be actively doing the science work in order for students to learn. (VDOE, 2018a).

Another VDOE staff member adds, "Since the dawn of time, students have asked questions about 'Why am I learning this?' 'Why do I need this?' This is an opportunity for teachers to apply this to real world situations, so that that eternal question can start to be answered for students." (VDOE, 2018a). In the video, they address a tool being developed and funds from a grant being used to ensure that application across the state is uniform and valid, despite local creation and grading of the assessments. Teachers at the professional development sessions geared toward developing performance assessments noted their support for this method saying things like:

We really want students to experience learning in relevant contexts, and so it makes sense to assess them in more relevant contexts. Getting feedback on your assessment, being able to revise it, is much more like what happens in the real world for students, or for people, when they're on the job, and so it makes a lot more sense for us to use that kind of assessment on an ongoing basis in our classrooms in Virginia. (VDOE, 2018b)

This feedback refers to not only the use of performance assessments as a summative assessment, but a more continual formative learning process within the classroom, which is a shift in perspective about how to measure and understand student learning.

2022-2023. While the Profile of a Virginia Graduate and alternative assessment policies remain in place, the policy intentions have changed at the state level. On the VDOE's website, the order of presentation under assessment resources reflects the VDOE's prioritization of SOL standardized types of assessment. The first resources presented are all SOL test related while the Profile information is presented second.

While the patron of the 2014 bill that initiated performance assessments in Virginia noted her confidence in the assessments as authentic, VDOE policies reflected a different line of thinking. The state BOE requires districts to develop a Balanced Assessment Plan (BAP) for submission to the VDOE only for courses offering performance assessments. This in and of itself sends a policy message to schools and school divisions about the assessments respected and trusted by state policymakers. In 2019, the BOE developed and VDOE disseminated to school divisions "Guidelines for Local Alternative Assessments for 2018-2019 through 2019-2020" (Virginia BOE, 2019). This document addressed and discussed the legislative decision that reduced the numbers of SOL tests given and directed that the courses no longer assessed by SOL tests would be required to assess students with alternative assessments.

By 2018, the state required the use of a performance assessment for verified credit in writing as addressed more in the accountability section below. This 2018 version of the guidelines document, and its follow-up, "Guidelines for Local Alternative Assessments: 2021-2022 and Beyond," (Virginia BOE, 2021) unmoored both assessments from the traditional school accountability system. In the 2021 version of the Guidelines, there was a reference to the

requirement of Balanced Assessment Plans (BAPs). While neither of the Guidelines documents

mention the *Profile of a Virginia Graduate* or the 5 C's, the guidelines for BAPs states:

> Detailed documentation of the daily formative assessment practices and strategies used
>
> during classroom instruction is not required, but the division's plan for supporting the
>
> effective use of ongoing formative assessment should be described. Evidence of
>
> opportunities for students to demonstrate acquisition of Virginia's 5 C's (critical thinking,
>
> creative thinking, communication, collaboration and citizenship) is encouraged. (VDOE,
>
> n.d.)

It is noteworthy that while the SOL tests and creation of a BAP are required, the evidence of the

5 C's is merely encouraged. This suggests the VDOE's first priority in 2021 "and beyond" has

shifted back toward content and away from an emphasis on the 5 C's as central to student

success.

Standards of Learning

Learning Loss and Growth. Discussions among BoE members and staff at BOE

meetings in 2021-2023 indicated a shift in the Board's focus back to Standards of Learning and

measuring student learning via standardized assessment. In the 2021-2022 school year, the

national conversation of "learning loss" dominated the education scene due to the impact of the

COVID-19 pandemic, especially related to school closures (Ferlazzo, 2021; Mervosh, 2022;

Turner, 2022). Virginia was no exception to this. In 2021, the Virginia legislature mandated that

Virginia Public Schools implement assessments for the purpose of measuring growth throughout

the school year, instead of just at the end. These assessments began in the Fall of 2021 in grades

3 through 8 in math and reading. These growth assessments are completely tied to the SOLs and

use existing SOL test questions. SOL tests, then, are now given in fall, winter, and spring, with

spring scores counting as a summative assessment, thus driving instruction. The 2022-2023 school year is the first year where all three tests are given using on-grade level content (Observation, BOE Work Session, February 1, 2023). While these tests cover different subjects than the current alternative assessments, it is possible that increasing the number of tests in math and reading will result in further narrowing of the curriculum, reducing the emphasis on science, social studies, and the "authentically" tested subjects. For these growth assessments, the VDOE provides teachers with "Growth Assessments Test Blueprints" that let teachers know which content areas will be covered on the tests. These blueprints reflect a state ideology that teaching to the test is normative and even encouraged. After each assessment, teachers are given student data by question so that they know which skills students are and are not mastering. Such state-sponsored increased emphasis on testing in math and reading will most likely increase the pressure teachers feel to drill the skills that students miss on the fall and winter assessments which will further lead to the narrowing of the curriculum.

Cut Scores. One significant debate within the BOE in the 2022-2023 year has been the politically motivated argument accusing the BOE of lowering standards for students. These accusations derive from the BOE's process of determining test cut scores. Cut scores are the scaled score cut offs that indicate the differences between passing and failing SOL tests. While cut scores are often framed as a technical assessment decision, they have also been maligned as the purposeful lowering of standards, often mentioned in the media and in Board meetings especially by recently appointed conservative VDOE administrators and BOE members. And since cut scores are perceived as a technical assessment reporting decision, unrelated to the standards that teachers use to teach within their classroom, this is a critical example of how political rhetoric doesn't always match policy language. To further illustrate the point, in the

CNN Town Hall in March 2023, the Virginia governor stated, "One of the big challenges, of course, is that…expectations and standards were lowered, for all Virginia kids. And sadly, they met those lowered expectations" (Youngkin, 2023). Despite the governor's assertion, there was never a decision made by the legislature, BOE, or VDOE to take an existing SOL test and lower the cut scores. BOE discussions around these cut scores have only occurred when new tests were implemented, and a decision had to be made to choose scores for proficiency (Observation, BOE Meeting, November 17, 2022). Cut scores were informed by expert review committees, and a Board member and staff person clarified that while there were changes made in the scores to determine proficiency, the discussion was not about lowering the standard for political purposes. Instead, when tests were updated, often time levels of rigor on the test were increased, and cut scores were changed to accommodate that shift in the assessments. In other words, while standards remained the same, the number of questions that a student needed to get correct in order to receive a scaled score of 400 and be identified as proficient were lowered.

During Board discussion about cut scores in November 2022, one longstanding Board member questioned the rhetoric surrounding the cut score debate and wanted to know the answer to the research question, "do higher cut scores actually lead to higher success?" (Observation, BOE Meeting, November 17, 2022). This question is really at the heart of the issue investigated in this CPA. While BOE members spent an hour and a half learning and debating about cut scores, is this decision mostly about optics? On an appointed Board, members often say they are not educational experts and defer to the hired education specialists and psychometricians for decision making. However, when press and political pressure are brought to bear, time is often spent in debate that may be unrelated to the impact on the student outcome goals the BOE rhetorically asserts they prioritize.

Resistance. Since the implementation of the SOL tests, parents have objected to their students' participation in SOL testing, especially at the elementary level. Shared on "Virginia is for Teachers," a teacher-run blog by a collaboration of teachers from around the state, one teacher shared that if parents refuse participation, their student's score will be reported as a zero for school accountability purposes (Sabiston, 2016). If 95% of students take the tests, the schools are not required to average in the zeros for reporting purposes, however, if more opt-out, it can have a negative impact within the school-level accountability system. Since student retention decisions are not permitted to be made on SOL scores alone, there is little consequence for the students who do not take the tests in elementary school at the individual-level; however, in secondary schools there are student-level accountability requirements tied to SOLs, such as verified credits for graduation, where students have to pass the course and the SOL in order for the class to count toward their graduation credits. Information from sources external to governmental institutions, like Virginia is for Teachers, is essential to understand the context of the policy because on the current VDOE webpage about the SOL program, there is no information about ways for students to opt-out, only a declaration, "All students in tested grade levels and courses are expected to participate in Virginia's assessment program" (VDOE, 2023a).

In 2021, news media reported there was an uptick in the opt-out numbers for SOL testing due to the additional stresses of the COVID-19 pandemic (Williams, 2021). In several 2022 BOE meeting discussions about accreditation, Board members mentioned in passing the percentages of students taking SOLs. There has been significant media coverage about the topic, with a range of reporting contesting the validity of the concern (Bass, 2022; Hunter, 2021; Nocera, 2022).

In addition to parental opt-out, there have been methods of resistance against standardized assessment systems enacted by teachers in less public ways. Hyde (2007) explains, "It is not that educators cannot or do not offer resistance in this era of intensified educational accountability but that, for the most part, educational researchers and teacher educators do a poor job of recognizing and naming their resistance" (p.v). Here, again, notions of power are important as teachers may be less likely to call attention to their resistance in order to best allow it to continue. Attention to resistance at the teacher-level would likely result in pressure to conform. Additionally, the various contexts and political cultures in which such resistance occurs is an important factor in the outcomes of resistance. Virginia is a "right to work" state, meaning that union organization is limited and constrained. In such a context, teachers have limited labor rights and diminished power in the state-led school structure, leading to less organized and effective resistance.

This does not mean that there are no ways for teachers to resist oppressive systems successfully, though. The Intercultural Development Research Association (IDRA; 2022) published "What Virginia's Anti-Equity Executive Order 1 and Reports Mean for K-12 Schools and Students: A Guide for School Leaders" that encourages teachers, administrators, and district leaders to resist policies that undermine equity. The discussion in the Guide is framed in the hope that, with activism, pathways for equity available through ESSA and resistance against SOLs and the current accountability system in Virginia can continue. Examples of resistance recommended in the Guide included continuing to advocate for unmet needs of students from marginalized backgrounds, documenting adjustments at the school-level that happen because of the executive order banning "inherently divisive concepts," reporting civil rights violations to the ACLU, and engaging in challenging the BOE policies through public comment (IDRA, 2022, pp. 24-25).

The authors of the report noted that these actions are necessary both to protect students and to help combat the "strict test-based accountability and expanded school 'choice,' that the administration will likely pursue in coming years" (IDRA, 2022, p. 21). It is the authors' hope that advocacy by local stakeholders can help promote continued flexibility and more nuanced systems of testing and accountability.

Considerations of New Assessment

While the Board's opinions and positions about standards and assessment have often been divergent and thus ambiguous based on the political affiliation of the governor who appointed them to the board, one Board member who was recently appointed by the Republican governor has made several recommendations to reconsider the entire system of assessment. In one such recommendation he stated:

> I believe at some point we have to conceptually decide what would we ultimately like to have and build backward for that, put it in RFP, and I think … [in] Virginia we [would] have lots of respondents to it.…This is big, big, big work, but this is exciting work, and I think it's the enduring legacy work.…In general, SOL tests, as we still largely know them, even though they're on computers now, even though they're using technology enhanced items, they're still the vestiges of the problem we solved in the 1990s. We solved that problem. We have new problems now and so we think about different blocks to use. (Observation, BOE Work Session, February 1, 2023).

At the beginning of the statement, when he advocates that the Board should "build it backward," he is advocating for the BOE to decide what its goals are and to build an assessment system, with standards and accountability, that will help students to reach those identified goals. As was shared in the previous quote by the former Superintendent of Public Instruction in the standards

section of this paper, the creation of the original standards and assessment system provided

continuity in instruction across the state and helped to identify schools that were struggling.

Since that has now been putatively accomplished, this BOE member cited above seems to want

to move away from the traditional SOL assessments and ways of measuring student learning that

have always been used. He indicates this in the last part of his quote when he mentions the

"blocks to use," a reference to an earlier BOE discussion about how test items are like blocks

that can be configured in different ways. So, in this statement, he is thus advocating for a new

system of assessment altogether, with shifts in standards and accountability to accompany it.

While he does not explicitly say what the new problems are, his intention in this statement is

clearly to consider innovating and changing the longstanding methods of measuring student and

school success. With the work of the HB585 Work Group, which is mandated to discuss new

systems of assessment, and with some members of the BOE advocating for new systems, there

are possibilities that sweeping changes could happen in upcoming years.

Both citizens and policymakers of Virginia also are thinking about system-level changes

in assessment. In the media review conducted between 2016 and 2023, nine columns and

editorials were published supporting changes to the assessment systems employed. Some, like

Delegate Coyner (2021) advocated for smaller changes like the inclusion of growth in

standardized assessment, which was ultimately adopted. Others advocated for portfolio-based

systems or more individualized plans (e.g., Alderson 2022; Mahoney, 2020). While there is not

significant consensus in the recommendations offered in these reports, the articles acknowledge

there are alternative systems of assessment and accountability that the authors find preferable to

the currently enacted SOL assessment system in Virginia policy.

Thinking about Equity: Policy Winners and Losers

While the overall national history of assessment makes it clear that policymakers should be considering racial equity when constructing and reforming assessment policy, in the Virginia Standards of Accreditation (SOA), there are no mentions of race related to assessment outside of inclusion of the federally required "subgroups" (Standards of Accreditation, 2018). All racial discourse within the BOE's SOA, VDOE website, and the State's federal accountability plan, is framed through the deficit lens explicitly called the "achievement gap." This suggests that the Virginia policy stance is one that is color-evasive, except when constructing a rhetoric of expectations that students of color need to perform better. While there is discourse about the need for scores to improve, there is little evidence of policy support to help students grow. Attention to considerations of the "achievement gap" in relation to the accountability and accreditation system are discussed in the following sections of this paper.

While often distractions from issues of equity are opaque and hidden in coded rationalized language, much of the distractions for racial equity by the current governor and VDOE administration have been overt and public. Prior to the election of the current governor, the VDOE had developed an equity division and an equity plan for the state. Equity was putatively a significant goal of the previous Governor (2018-2022), under whose leadership the VDOE created a new department, the Office of Equity and Community Engagement, situated within the newly renamed Department of Policy, Equity, and Communications. After the 2022 election, all mention of equity was scrubbed from the VDOE website and its initiatives. The Virginia Office of Equity was disbanded and the department that housed its name reverted to the Department of Policy and Communications.

The current conservative governor is reported to have stated that equity is a "very confusing word" (Natanson & Vozzella, 2022) and is claiming that "we shouldn't embrace equity at the expense of excellence," (Youngkin, 2023). Excellence implies the same level of program delivery for all members of the educational system despite unequal conditions while equity frames the issue in terms of attending to historical discrimination of some members in the educational system. Equity and excellence oftentimes conflict as it does in the standards, assessment, and accountability system in Virginia. The Governor's comment is an example of how the values of policy actors inform policy decisions. The idea of embracing equity at the expense of excellence indicates a belief that assessment policy should attend to excellence for all over focusing on equity for marginalized students. Here, competing values from American culture, quality and equity (Marshall et al., 1989), are in tension, but excellence is the primary value promoted by the VDOE, as the first component of their Code of Ethics states, "As employees of the Virginia Department of Education we will dedicate our efforts toward excellence in public education through continuous improvement" (VDOE, 2023c).

Accountability and Accreditation: Ambiguities in Policy Discourse

Standards and assessment policies are intimately connected to accountability policy. In Virginia's current system of accountability, it is evident that ambiguities in policy discourse lead to mixed messages for district administrators, building administrators, and classroom teachers (see papers two and three). In many states across the United States, accountability and accreditation are constructed differently. In the Virginia system, state accountability and accreditation are part of one system, but the federal accreditation mandate has different reporting requirements about achievement and staffing than the state system.

In terms of the competing or divergent goals reflected in the state standards, assessment, and accountability systems, while schools in Virginia are expected to develop BAPs and there are verified credits that can be achieved using authentic performance assessments, the VDOE website for accreditation and accountability only refers to the Standards of Learning, with no mentions of the Profile or the 5 C's. Additionally, within the accountability plan submitted by the VDOE for federal accountability, there is no mention of Profile or the 5 C's policies. These are indications of policy priorities and ambivalence on the part of policymakers to fully embrace the new Profile of a Virginia Graduate and its consequent performance assessment approach. In making decisions around accountability and accreditation policies, interactions and negotiations within BOE meetings reveal the policy interpretations made by Board members and just how much partisan politics influence decision making processes and policy outcomes. In Virginia, assessment policy is political.

Original Accreditation Policy Intentions, 2017-2018

As previously mentioned, in the 2018-2019 school year, changes in accreditation policies and procedures were implemented by the BOE that allowed for both proficiency and growth to count toward school accreditation. In the introductory messaging of the changes by the VDOE, the Assistant Superintendent for Student Assessment and School Improvement noted why these changes were conceived and implemented:

> We have had an accountability system in Virginia for almost 20 years now and we've certainly learned a great deal about accountability during that time period. So, I think one of the goals of the new accountability system is to look at what we learned from that previous system and to try to make some improvements….One thing we learned with the previous system was that there was concern that, because we were looking just at pass

rates, whether or not students were proficient on a test, we weren't capturing the progress

of students who perhaps didn't pass the test but had made a great deal of growth during

the year. The previous system didn't account for that progress that students had made, and

so in the new system we will attempt to more accurately capture the growth that students

show. (VDOE, 2017b)

As stated, the purpose behind the new practice of including growth in considerations of SOL

pass rates was that schools where students were starting at a lower baseline were not punished,

and when schools were able to show growth for those students, that was, and still is, counted as

an indicator of progress worthy of credit toward school accreditation. The way this was

accomplished was that students who did not pass the SOL but made a growth of a certain amount

from the previous test, were included in a combined pass rate. Additionally, English language

learners who did not pass reading tests but made advancement on the English Language

Proficiency test were also counted towards the combined pass rate for the school's accreditation.

Proponents of this model of accreditation praise it for its careful consideration of students from

disadvantaged backgrounds, English language learners, and special education students who are

often invisible, especially in wealthier districts, when overall proficiency scores are all that is

reported, and their scores do not threaten accreditation.

The decision was also made by the BOE during this period to allow for either the single

year combined pass rate, or a combined rate of the previous three years to be used for

accreditation reporting, whichever was higher. As one Board member summarized recently, "the

three-year trial was intended for high performing schools to be able to innovate" (Observation,

BOE Meeting, September 14, 2022), so that if a new practice that was implemented didn't go

well, there would be less disastrous consequences for the school. Teachers indicate that this type of flexibility is necessary if any growth is to happen in Virginia school systems (see paper two).

Additionally, there was a slight shift in the stakes associated with the new Standards of Accreditation (SOA). While *School Quality Report Cards* would still be published, the new system did not label any school as "failing." Instead, the majority of schools were either accredited or accredited with conditions. A past Board member describes the rationale behind this decision:

> Our system made it very difficult for schools that were identified as, and I'm going to use a term that's used out of the public, it's not one that I would like to use but 'failing schools' that are not meeting the needs of all their students. Schools that would not be fully accredited, it is hard to get out of that designation, and once that designation comes into place, teachers don't want to teach in a 'failing school,' so to speak...we've created a system that disincentivizes folks to be in that school. (VDOE, 2017b)

In the new system, to avoid the 'failing' label, schools would be accredited or accredited with conditions, which would require a submission of a school improvement plan. The only way for a school to not be accredited was if they were not following the school improvement plan they created after being designated as "accredited with conditions." The idea behind this change was to lower the stakes of the accountability policy and make a shift from a punitive system to a supportive system. As discussed by the Board member above, the hope was that this would allow for more support at schools that were struggling to meet the required proficiency targets, especially in regard to teacher retention.

Pandemic

During the 2020-2021 school year, many schools were accredited without meeting the markers of the system due to the nature of the pandemic. One of the markers of the system was chronic absenteeism, and COVID's required quarantining impacted this indicator significantly. Additionally, decisions were made by the BOE to cancel SOL tests while schools were closed in 2020, and participation in state testing dropped in 2021. This had a negative impact on the three-year test score averages option built into the system referenced above.

Additionally, across the country, students scored lower on the NAEP in the 2021-2022 school year than in previous years. In the discussion by BOE members, VDOE staff, and in the governor's media releases, there was no mention that what students should learn by a certain age or grade is a social construct, and that allowing grace after multiple years of social and academic disruption should be taken under consideration. Instead, students were expected to maintain grade-level proficiencies and compare at rates matching students from the previous years who had not experienced a global pandemic. While many adults were allowed flexibility in their levels of productivity while adjusting to work from home, this allowance was not considered for students, who were hounded for their learning loss and drops in SOL and NAEP scores. Some of the negative attention in response to scores was likely due to the pressures of meeting benchmarks of accountability that have, in many ways, replaced the goal of authentic student learning. This is seen in the current policy conflict surrounding student proficiency and accountability.

Accountability Policy Conflict: 2022-2023

Interestingly, a very similar rhetoric to the one used when the new SOA were first implemented in 2018 is being employed currently, but from the counter perspective. While in

2017, VDOE publications explained the concern was that growth was masked, in 2022 members of the BOE, VDOE, and the governor have displayed concern about masked rates of proficiency:

> "[The Superintendent of Public Instruction appointed by the conservative governor] in an interview said growth and proficiency should not be lumped together because it does not adequately measure the learning loss that students are facing across Virginia. She pointed to state assessments and national test scores that reflect historic learning losses in core subjects in math and reading, arguing that the data evidently shows students have substantially struggled." (Vozzella & Asbury, 2022, paragraph 29)

So while accountability requires distilling many data points to make accreditation decisions, there is disagreement by state policy actors over which data points should matter most and how to best manage and report that data. This is yet another example of the role of politics in the policy process.

The current conservative governor has echoed these critical remarks about accreditation, especially related to COVID learning loss. While the governor finds the accountability system "broken," those who still support the current SOA assert that the governor "manufactured that there's this crisis in public education" (Vozzella & Asbury, 2022, paragraphs 3 & 15). This idea of manufacturing public urgency around the state of schools is exactly what was accomplished in the politics of *A Nation at Risk*, which Berliner and Biddle (1997) also referred to as a "manufactured crisis." Some current Board members are harkening back to that rhetoric of crisis in current discussions around accreditation policy. For instance, one recently conservatively appointed Board member, who was on staff in the U.S. Department of Education when *A Nation*

at Risk was released, said that he fears we are again part of a "rising tide of mediocrity" (NCEE, 1987 as cited by Board member, Observation, BOE Work Session, February 1, 2023).

In consideration of updating the accountability system in Virginia, a consultant from the education foundation founded by former Florida Governor Bush was asked to present at a Virginia BOE meeting in October 2022. The choice of this presenter to appear before the BOE is in itself an exercise in power. Her position with an education foundation associated with former Governor Bush and the Bush political dynasty provides her with an aura of legitimacy and thus as a resource in the political debate for reaffirming traditional high stakes accountability systems. In this meeting, she encouraged the BOE to use the rhetoric of manufactured crisis. The presenter, a consultant for the construction of the accountability systems in Florida and Mississippi, repeatedly referred to the need to *create* a sense of urgency about assessment results in order to build policies for traditional high stakes accountability systems. For example:

> When you cite that three-fourths of your students are already proficient or advanced and are ready to do grade level or college level work, may that be, there's not a sense of urgency to act. You're not really inspiring anything to happen when you're saying that all your students are faring okay. (Observation, BOE Work Session, October 19, 2022)

This relates to the prior conversation by the BOE about assessment cut scores. The consultant was not commenting about whether the students are or are not proficient, she was instead making an argument that policymakers should change the metric. In the school systems she consults for, accountability grades are published as an A-F for each school, a decision previously reversed in Virginia policy. She noted that this was not a politically popular decision, but she believed that this is the cause of Mississippi and Florida's increase in NAEP scores. In reflecting upon history, the last time that states across the nation decided to employ a new system based on the growth of

NAEP from two states, with the implementation of standardized testing, the results did not generalize.

One of the major claims of that presentation to the BOE was that the graded accountability reporting system would help historically marginalized populations. With the creation of urgency within the context of "divisive" concepts rhetoric, this may well be a prime example of how the "creation of the policy problem [is] a distraction from matters of racial equity" (Bradbury, 2020, p.247). This is further explored in the next section which considers the relationship between the accountability system policy and historically marginalized student groups.

Achievement Gaps

The only mentions of race within any of the Virginia accountability policy documents reviewed were related to achievement gaps, the outdated term still used in the policies that normalizes a deficit model of policy thinking within the standards, assessment, and accountability policy discourse (Stein, 2004). While, again, this is presented as concern for students with below-level achievement, there has been no effort to consider or employ the term education debt, opportunity gap, or others considered more appropriate in the literature related to assessment (Ladson-Billings, 2006; Weiner & Carter, 2013). While the Virginia Office of Equity was in place, there was a focus on equitable outcomes and resources distributed to assist with that goal. The 2022 VDOE administration chose to focus on "equal opportunities" instead, a color-evasive, one-size-fits-all approach (Leonor, 2022).

The current president of the BOE also prefers a color-evasive approach. He noted that it's a "pet peeve" that "everyone" refers to these gaps as the "Black-White Achievement Gap," noting that he thinks it makes more sense to contextualize the gap in terms of socio-economic

status (Observation, BOE Work Session, October 19, 2022). He further supported his preference with anecdotal evidence about discussing this with a researcher who asserted that there is better data on race than socio-economic status, which often means that data gets reported by race. At a different Board meeting, another Board member indicated that they felt similarly, stating, "I can tell you who's going to flunk those tests right now. We don't even have to give the tests, I can give you the economic data on Virginia, that will tell you the outcome" (Observation, BOE Work Session, February 1, 2023). By this, she was critiquing the sole reliance of the state on proficiency scores as a metric of student outcomes. She cited research that supported her claims that standardized test outcomes largely trend with economic data, indicating her understanding that standardized testing itself measures socio-economic status more than the desired learning outcomes and an admission that the assessments are sensitive to race and social class.

A more recently appointed Board member contested this: "I just have to say the demographic destiny talk makes me extremely uncomfortable. There are schools that beat the odds" (Observation, BOE Work Session, February 1, 2023). Here, he was indicating his belief that assumptions should not be made about who will and who will not score well on the tests based on the socio-economic factors; however, the term "beat the odds" indicates that he realizes that this is not typical and that many students, especially in economically disadvantaged neighborhoods, will be labeled as failing by the accountability system.

In a related Board meeting during which Board members were discussing the possibility of reimposing the A-F accountability rating, the consultant who had worked on the Florida and Mississippi accountability systems noted that while additional funds weren't given to schools that were rated D or F, they ended up getting more money because of the funding allocations of Title I since Title I funds are given to schools with a high population of students from

disadvantaged socio-economic backgrounds. Her comments indicated that the accountability metrics don't punish schools into achieving higher scores. Schools with disadvantaged students still have a difficult time meeting the accountability standards, even when publicly shamed by low school report card grades. While some schools may "beat the odds," a synecdoche, this exchange during a Board meeting provided further evidence to the idea that accountability systems relying on high stakes standardized testing may be measuring the economic composition of neighborhoods rather than learning outcomes.

If standardized tests are perceived as valid and normalized, and the goal of the accreditation system has been stated to "close gaps," then the system is failing. Recent NAEP reports indicated that gaps between marginalized students and their peers are not closing, and that this was especially true throughout the pandemic (Bryson, 2022a). The measuring of student standardized achievement outcomes by race and language status, without additional resources and support, is not producing student growth. While the policy goal of tracking achievement by race may be asserted by policymakers as a way to promote equity, without support and resources for students and teachers, such a system is merely punitive, and the outcomes of the policy do not match the putative goal. In fact, such policies reinforce deficit-thinking and continued marginalization of historically disadvantaged populations.

Resistance Around Accountability Systems

There is resistance in response to reverses on equity policies related to the Virginia accountability system. While the Virginia Office of Equity in the late 2010s and early 2020s had created policies that focused on helping schools and districts across the state achieve equitable outcomes, under the 2022 administration, the governor and VDOE dismantled the equity policy and replaced it with a focus on equal opportunity. A letter delivered to the Superintendent of

Public Instruction from the Association of School Superintendents stated that the VDOE's focus on equal opportunity over equitable outcomes does not allow consideration for the factors impacting many historically marginalized students and "set[s] public education in Virginia back many years (Leonor, 2022, paragraph 14). Further, they emphasized, "Quality education in Virginia has to be more than providing opportunities and hoping for the best. Virginia's accountability system relies heavily on student outcomes, not opportunities" (Leonor, 2022, paragraph 16). While the superintendents don't go so far as to advocate for changing the accountability system, they do indicate the need for resources to match the demands put upon schools by the VDOE's accountability policies.

Others, however, do advocate for changing the system of accountability. One school division board member said in 2018, "I think that we have believed for many, many years that focusing on accountability would fix things…We haven't fixed things; we've made a bigger mess" (quoted in Mattingly, 2018). The "bigger mess" is likely in reference to a cheating scandal that had occurred within her district in the last SOL testing window. It was discovered that amidst the high pressure for school achievement, many teachers at a particular school were assisting students on their standardized tests. While teachers and administrators were fired due to this decision, this school board member advocates for placing some of the blame on the system itself and advocates for the consideration of changes that might help "fix things."

Performance Assessments for Accountability

When the legislation to cut-back on the number of SOLs tests given to students passed in 2014, local school divisions were required to develop local performance assessments for measuring learning in science and social studies tests in Grades 3 and 5 as well as writing in Grade 5. However, while performance assessments data was reported to the state, it was

excluded as an accountability or accreditation factor. Many teachers interpreted this policy

change as an indication that reading and math, still assessed with SOLs and tied to

accountability, were more important than science and social studies, thus narrowing the

curriculum and their pedagogical practices (e.g., teaching to the test) to match that interpretation

of policy (see paper two). Additionally, by signaling to teachers that the alternative assessment

plans were not factored into the accountability system by explicitly stating: "The results of the

local alternative assessments will not be used to designate state accreditation or federal

accountability status," (Virginia Board of Education 2019, 2021), this message about the

primacy of the SOL tests was further amplified to school district, school administrators, and

teachers.

Starting in 2018, The BOE allowed school districts who followed the procedural

processes laid out by the VDOE to grant verified credits in writing and history for student-level

accountability. While this shift allowed for some student-level accountability that is less rigid

than standardized testing, the BOE has long been hesitant to tie accountability practices to

authentic or performance assessment. For instance, at the October 2019 Virginia BOE meeting,

Board members discussed their hesitancy in connecting a system of accountability with the

framework for capstone presentations, a performance-oriented approach, even though significant

time had been spent around developing protocols and support for districts that wanted to use

capstones (Observation, Virginia BOE Meeting, October 17, 2019). This trepidation indicated

that even with a willingness to explore alternative ways for students to show mastery of

knowledge, policymakers feel that the standardized assessments are more valid for assessing

mastery of content. The policymakers at the legislature, BOE, and VDOE have decided to make

nods to both standardized content assessment as well as performance assessment centered on

measuring alternative 21st century skills as reflected in the Profile and 5C's frameworks. This is a clear example of Rein's (1976) notion of intentional ambiguity in policy formation, so that consensus can be met. While the ambiguity allows those with differing perspectives to agree to the new policies, it leads to confusion at the implementation level. It is notable, however, that there are long-standing examples from New York and newer systems emerging in New Mexico and Colorado, in which schools have implemented locally constructed performance assessments in their accountability systems

Summary of Results: Standards, Assessment, and Accountability Policies in Virginia

Within standards policy, the introduction of the Profile of a Virginia Graduate in 2017 indicated a major shift from the normative content-based standards that had been in place since the mid-1990s. While the SOLs standards, assessment, and accountability system continued to operate, the Profile added on the higher-order 5 C's: critical thinking, creative thinking, communication, citizenship, and collaboration skills, as essential knowledge for Virginia students. Additionally, the review of the History & Social Studies Standards in 2021-2023 indicated at the state level there is mixed messaging in the policy about which of these approaches to educational outcomes, discrete content-based (SOL assessments) or higher order skill-based (authentic performance-based assessments), is most valued at the state level. In 2021-2022, the VDOE led a rewrite of the standards that emphasized inquiry-based learning with a focus on the 5 C's. When submitted for review to the BOE at the end of 2022, after a change in political landscape from more progressive to more conservative, the new appointees to the Board and VDOE rejected this shift toward higher order learning standards and performance assessment. New standards were quickly written that were more aligned with the traditional normative SOL standards approaches.

In terms of assessment, mixed messaging by policymakers about SOL tests and performance assessments have continued. While the original intention for reducing the number of SOL tests required was to give more space for the 5 C's, actual prioritization by policymakers regarding assessment has remained aligned with the standardized SOL tests. Rhetoric in conversations about the lowering of standards through cut scores were informative. Following the pandemic, more standardized testing was implemented with SOL tests administered in the fall, winter, and spring to allow for growth measurement, while summative assessments have remained tied to only spring assessment scores. Reminiscent of Stein's (2004) emphasis on policy culture and symbolic language, Virginia policymakers rhetoric around equity were highlighted in assessment policy with a focus on the "achievement gap" which views historically marginalized students through a deficit lens. This cultural deficit thinking was reflected in Virginia's assessment policy.

Finally, Virginia policymakers' emphasis on the importance of growth versus the importance of proficiency in the formation of accountability and accreditation policy in the system reflected conflicting policy values and produced mixed messaging. While some Board and VDOE members indicated that growth for students was an important indicator of school success and should remain in the school accreditation model, others noted that the inclusion of growth masked proficiency and gave false impressions of student and school success. In recent debates among policymakers in which conflicting goals have been discussed, signals such as the urgency for reconsideration of implementing an A-F grading scale in the accountability system indicate waning support for higher order learning approaches reflected in the Profile, 5 C's, and performance assessment approach. The interpretation of this policy ambiguity by educators at the school district and school levels are the subject for the last two papers in this dissertation.

Discussion

Value-Centered Critical Policy Analysis

Rein (1976) asserted that policy analysis should consider values since policy-formation is based on values. This CPA utilized Rein's (1976) guidelines for value-centered policy analysis, Young and Diem's (2017) tenets of Critical Policy Analysis, as well as questions from Bradbury's (2020) race-centered CPA, Guba's (1984) definitions of policy, and Flyvbjerg's (2001) guidelines to making social science matter to employ a value-centered CPA (See Table 10 for summary of key components of each). The results of this research centered on these concepts and the following discussion centers on those ideas as applied to the existing literature on education assessment and accountability policy (Brunn-Bevel & Byrd, 2015; Labaree, 1997; Marshall et al., 1989; Spillane, 2005; Stein, 2004) are discussed below.

Largely, the history of standards, assessment, and accountability policy in Virginia should inspire skepticism about the policy system as it currently exists. Add to that the consideration of shifting policy outcomes to ambiguous purported policy goals, and it becomes clear that mixed messaging from the policy-in-intention level abounds. Major reasons for these mixed messages are an unwillingness to talk about race and power, other than through a deficit-lens, and the role of degenerative politics in the policy process. There are policy advocates, teachers, students, and policymakers who are challenging the status quo, and this is a hopeful sign for change. The discussion will conclude with considerations about viable next steps.

Table 10

Theoretical Groundings Informing the CPA Discussion

Combined Discussion Topic within this CPA	Rein (1976)	Young and Diem (2017, p.4) Concern regarding:	Additional Conceptual Guides
Policy Origin	"Examine Policy in Historical Perspective" (p.23)	"the policy, its roots, and its development (e.g., how it emerged, what problems it was intended to solve, how it changed and developed over time, and its role in reinforcing the dominant culture)"	"Is the creation of the policy problem a distraction from matters of racial equity?" (Bradbury, 2020, p.247)
Rhetoric versus Experience	"Attend to Practice as Well as Policy" (p.22)	"the difference between policy rhetoric and practiced reality"	"What is the impact of the policy on pedagogy and practice? How does the policy produce practices that result in disparities in attainment through seemingly neutral practices?" (Bradbury, 2020, p.247) "How does this policy present those in power as caring about racial equity (even where this is not the result)?" (Bradbury, 2020, p.247) Policy-in-intention, policy-in-action, policy-in-experience (Guba, 1984)
Mixed Messaging	"Treat the Question of Purpose as Unresolved" (p.20)		
Issues of Power, Race, and Privilege		"the distribution of power, resources, and knowledge as well as the creation of policy 'winners' and 'losers'" "social stratification and the broader effect a given policy has on relationships of inequality and privilege"	"How does this disadvantage one group more than another? What is the purpose of this?" (Bradbury, 2020, p.247) "How does the absence or presence of 'race' perpetuate inequalities?" (Bradbury, 2020, p.247) "Is the creation of the policy problem a distraction from matters of racial equity?" (Bradbury, 2020, p.247) Culture of Education Policy (Stein, 2004)
Degenerative Politics	"Consider the Political Reception of Policy Studies" (p.30)		Political Spectacle (Edelman, 1988, as cited in Smith et al., 2004) Degenerative Politics (Ingram & Schneider, 2006)
Challenging the Status Quo	"Distrust Orthodoxy" (p.25)	"the nature of resistance to or engagement in policy by members of nondominant groups"	
Role as a Moral Critic	"Approach Social Policy as Moral Critic" (p.26)		"How does policy constitute groups or individuals – as problems or part of solutions?" (Bradbury, 2020, p.247) "What should we do?" (Flyvbjerg, 2001, pp.136-7)

Policy Origin

Critical perspectives on policy analyses emphasize the importance of the historical

context and origin of the policy (Bradbury, 2020; Rein, 1987; Young & Diem, 2017). The origin

and history of the standards and assessment policy in Virginia give reason to be concerned about

over-reliance on the traditional system of standards, assessment, and accountability reflected by a

continuing emphasis on the SOLs. I have attempted to illuminate the relationship between racism

and standards, assessment, and accountability systems in Virginia. While there is a history of

considering education the great equalizer, an education system based on racist histories will

struggle to meet that aspiration (Au, 2016; Gould, 1996; Randall, 2021, Randall et al., 2022).

Additionally, the deficit-mindset from which the assessment and accountability policy arose

continues on through implementation phases of the policy-in-action and -experience and

encourages those with deficit-mindsets about their students to perpetuate those stereotypes

(Stein, 2005).

Policy Rhetoric Versus Experience

Guba (1984) clearly delineated that policy differs as it moves through levels from policy-

in-intention, to policy-in-action, to policy-in-experience. With that understanding, Young and

Diem (2017) and Rein (1976) asserted that differences in policy and practice must be explored.

This CPA determined rhetoric about putative beneficiaries and policy outcomes are unaligned,

which is further explored in papers two and three.

Assessment policy is one of the primary mechanisms through which values are translated

into policy (Marshall et al., 1989). As seen in this CPA of the Virginia standards, assessment,

and accountability system as a whole, there are often differences between the policymakers'

ideologies, political standpoints, and rhetoric that have an impact on how they create and

implement policy mandates. One addition of this study to the greater literature is a deeper understanding of how policy-in-intention and policy-in-action, as defined by Guba (1984), are influenced by the political culture, structure, and politics within a state. It illuminated how multiple and conflicting values lead to ambiguous policy mandates with mixed messages for practitioners. At the end of the day, it is clear that Virginia policymakers are ambivalent and uncertain about the goals of the SOL system. They have entertained the idea that the curricular standards, bolstered by the assessment and accountability systems are problematic and therefore Virginia public schools should be shifting to a focus on 21st century higher order learning. Unfortunately, they appear stymied by their inability to link performance assessment to an accountability system. They appear to be teetering on the edge of reform but stifled by partisan political pendulums shifting back and forth, in a state of arrested development.

Mixed Messaging

The purposes of policies often change, even when the policy itself does not (Rein, 1976). This results in mixed messaging across all of Guba's (1984) definitions of policy. While the policy itself hasn't changed, the meaning that policy actors bring to it, as well as implementers, are not stable. This was illustrated clearly in the competing SOL standardized assessment and the 5 C's centered performance assessments. The 2017 BOE noted that the purpose of the SOLs had changed since their earliest conceptions because in the mid-2010s schools were overly test-focused to the deficit of learners. This was the impetus for the introduction of the Profile of a Virginia Graduate, which was to allow space to focus on the 5 C's. The 2023 Board has since retrenched this policy stance and refocused their attention on standardized test scores despite the fact that the Profile and accompanying policies remain in place.

Purposes of Public Education. A clear example of mixed messaging can be seen in policy documents referring to the purpose of the Virginia Public Schools, which are summarized in Table 11.

Table 11

Virginia State Policy Actors Perspectives on Purposes of Public Education

Policy Actor	Quote	Purpose(s)	Value(s)
Enacted Law by Legislators & Governor in 2016	"a Profile of a Virginia Graduate that identifies the knowledge and skills that students should attain during high school in order **to be successful contributors to the economy of the Commonwealth**, giving due consideration to critical thinking, creative thinking, collaboration, communication, and *citizenship*" (SOQ, 2016)	Primary: Training workers or economic mobility Secondary: Democratic citizens (Labaree, 1997)	Efficiency, excellence, & choice (Marshall et al., 1989)
Board of Education	Priority number 1 (2018-2023) "provide high-quality, effective learning environments for all students" (VDOE, 2023b). SOA (2018): "The **foremost purpose** of public education in Virginia is to provide children with a high quality education **giving them opportunities to meet their fullest potential in life**….The mission of the public education system is to educate students in the essential knowledge and skills in order that they may be *equipped for citizenship*, *work*, and an informed and successful life."	Social mobility (Labaree, 1997) Lifelong learners	Equity (Marshall et al., 1989) Holistic development
Virginia Department of Education (2018-2022)	Creation of the Office of Equity and Community Engagement (Disbanded in 2022)	Social mobility (Labaree, 1997)	Equity (Marshall et al., 1989)
Code of Ethics	"As employees of the Virginia Department of Education we will dedicate our efforts toward **excellence in public education** through continuous improvement (VDOE, 2023c)	To be the best	Excellence
Governor	"we shouldn't embrace equity at the expense of excellence" (Youngkin, 2023)	To be the best	Excellence

Note. Emphasis added in all quotes.

While all of these definitions occur at Guba's "policy-in-intention" level, Table 11 indicates the diverse purposes to which public schools are held (Labaree, 1997) and aligned values (Marshall et al., 1989), which contribute to mixed messaging both among policy makers and policy implementers. The primary purpose of policies broadly conceived, and assessment policies specifically, in the eyes of the legislature is centered around economic production, which could be viewed as the training of laborers or as a preference of economic mobility. This provides an umbrella for the unclear purposes and vehicle for mixed messaging in policy. In the legislative policy, there is a secondary reference to preparing democratic citizens. These purposes align with the values of choice, excellence, and efficiency, in which all students will be trained as citizens in efficient ways that allow them to make choices about their own interests in the future.

For the BOE, the primary purpose of education is outside of the frameworks provided by Labaree (1997) and Marshall et al. (1989), with a focus on lifelong learning for holistic development. However, within the purpose of schooling written into policy documents, they also emphasized preparing citizens and workers. In the written Board priorities, they note the desire to "provide high-quality, effective learning environments for all students," a nod to the value of equity. Together with their declaration of the purpose of schooling, it follows that there is some belief in school as an equalizer that can promote social mobility. While this was also a priority at the VDOE prior to 2022, the Board has shifted the main value toward excellence, as noted in their Code of Ethics and by the current governor's remarks that "we shouldn't embrace equity at the expense of excellence" (Youngkin, 2023). So, the current degenerative political forces are pushing excellence as a primary value, in contradiction to the purposes of the public schools as asserted by the Standards of Accreditation and the priorities of the Board.

Marshall and colleagues (1989) note that policy makers have long struggled to embody equity in policy, as it is presented as a founding ideal (e.g., "All men are created equal;" p.91). Excellence necessitates hierarchy, and when excellence is chosen at the expense of equity, students who have long been disadvantaged by systems of accountability fall further behind (Heinecke et al., 2003). Labaree (1997) argues that conflicting goals are not necessarily problematic, as graduates of public schools, in fact, do need to be taxpayers, workers, and consumers. However, the results of this CPA indicate that mixed messaging about these diverse goals creates ambiguity about the priorities of the education system and assessment policy. Some policy goals converge and some diverge. If it is true that all of these goals are equally important, components of them need to be built into the assessment and accountability system. If there are values and purposes that are secondary, policymakers need to be aware of how messaging around that will impact interpretation by educators who are responsible for implementation. It is clear that those purposes and values not supported by specified means within the accountability system will be given little attention by those implementing assessment policy.

The Process of Mixed Messaging. One conception of mixed messaging is that as policy is translated through the levels of implementation, its intention is lost little-by-little or distorted like a game of telephone (e.g., Spillane, 2005). The results of this CPA, conceptualize mixed messaging in a different way. Intentions don't simply distort across levels of implementation (e.g., policy-as-intended, to policy-in-action, to policy-in-experience; Guba, 1984), the intention is also ambiguous at the initial policy-as-intended level.

Ambiguous policy is often used as a tactic for consensus (Rein, 1976). While this creates tension in lower-levels of implementation, it also creates tension during policy revisions and evaluations during policy formulation. This was exemplified in the History and Social Studies

Standards update when the ambiguity between the directives of the Profile and the SOLs played out at the Board-level with divergent perspectives around what the updated standards should look like.

The Profile of a Virginia Graduate was also an example of this. Smith and colleagues (2004) assert, "When a policy lacks an instrument or provides an instrument so weak that it could not reasonably be expected to effect the desired outcome, that policy falls into the category of hortatory or symbolic" (p.6). Because the measures of accountability remained tied to the SOL tests rather than the performance tasks and the 5 C's, the instrument to mobilize the policy goals became so weak that the 5 C's were regarded as symbolic, as discussed in more detail in paper two. With this accountability policy, without purpose and intention clearly explained, it is not just implementers who struggle with the purposes of policies, but also Board members. In other words, the generation of ambiguity is not only in the translation of the policy from legislature, to BOE, to VDOE, but the interactions within each of these bodies, with their own agency intentions that are making the policy outcomes unclear. This becomes even more accurate when politics are considered, as discussed in the below section on degenerative politics.

Issues of Power, Race, and Privilege

In order for a policy analysis to be considered critical, notions of power must be centered. Young and Diem (2017) encourage considerations of power in terms of who the "winners" and "losers" of any policy may be, with an added focus on social stratification and how the policy may be recreating systems of inequity and privilege (p.4). It is interesting to consider who the real winners of the current Virginia standards, assessment, and accountability policies are. Papers two and three help illuminate that teachers do not see themselves or their students as the winners in this policy game.

The system does, however, provide advantages to those who are already fiscally advantaged. Bradbury (2020) asks, "How does this [policy] disadvantage one group more than another? What is the purpose of this?" (p.247). Students of wealthy citizens are winners within the Virginia assessment policy system because they have the financial wherewithal to circumvent the barriers of the system. As Ingram and Schneider (2006) put it, many of our policies result in "an American democracy that espouses ideals of equal protection and treatment under the law, while actual treatment by policy of citizens is noticeably and unfairly unequal" (p.171). When students whose parents have means do poorly on tests, they can get tutors. Students of wealthy families live in neighborhoods where taxes for neighborhood schools are higher and therefore the schools have more resources and attract better teachers who are less likely to face burnout. Histories of racism, redlining, and elitism in Virginia speak to original purposes that relegated inferior resources to students of color and economically disadvantaged students. Even if there has been progress in societal views with regards to race and class, when policies remain face-neutral, they result in widely disparate outcomes, in large part due to the fact that there has never been redress for the racist and classist histories of the systems of education.

Stein (2004) argues that the labeling in policy language of certain students as inferior and needing saving is part of the policy problem. While she focuses on Title I and the labeling of schools as poor and deficient, the same is seen in the debate around students who are labeled below proficient in assessment policy language. While basic reading and math skills are important, as one Board member pointed out, there is not sufficient support that the changing of cut scores or these labels within an accountability system result in improved outcomes for students.

This focus on cut scores, proficiency rates, and excellence, then, is political spectacle. In Bradbury's (2020) CPA, she asks "How does policy constitute groups or individuals – as problems or part of solutions?" (p.247). As Stein (2004) notes that labeling students as deficit reifies the status quo, this labeling of students as deficient and not excellent is the creation of a policy problem rather than an eye toward an education solution. While the raising of cut scores and promotion of excellence seems neutral, history shows that these policies result in continued advantage to wealthy, White students without a commitment to a solution for students of color or economically disadvantaged students.

Assessment and accountability are especially implicated in production and reproduction of social stratification, with notions such as "achievement gaps" that center the academic performance of marginalized students without the necessary support to provide pathways to success. While transformation to the term "opportunity gap" (Weiner & Carter, 2013) or "educational debt" (Ladson-Billings, 2006) has the potential to be a change in symbolic terminology rather than authentic shifts in mindsets, beliefs, and systems, the lack of such thinking in Virginia policy indicates a policy culture trapped in a culture of deficit language as reflected by continued framing around the achievement gap and its damaging potential. This description of students as deficient is not new, and this conception of students within Virginia's policy in intention translates to and encourages deficit-mindsets at the policy-in-action level (Stein, 2004). While some may consider this reproduction of social stratification unintentional, the recent dismantling of the Virginia Office of Equity, white-washing of social studies standards, and push to remove growth as a factor of accreditation, indicate that the policy decisions encouraging social stratification through assessment policy in Virginia are likely intentional. A reason for this intentionality can be attributed to the political culture of Virginia as

well as the fact that politicians often become the winners of policy, as discussed in the next section on degenerative politics.

Degenerative Politics

Value-centered policy analysis acknowledges the political nature of policy formulation and implementation, as well as the fact that policy research may be used in ways that were unintended by the analysts (Rein, 1976). As Smith and colleagues (2004) put it, "policies flow out of politics and…politics flow out of policies" (p. 2). This is especially true in the politics surrounding assessment and accountability. This is detrimental to the students who should be served by education policy.

Within this analysis, the power of political culture of a state and the political context of policymaking have been shown to have significant impacts on policy formulation and implementation. Politicians are using education policy to further their own partisan agendas. This was evidenced by the political appointments of BOE members and the Virginia Superintendent of Public Instruction who leads the VDOE, as well as political influence over public opinion. This is not a tactic only employed by one party, as both parties try to motivate their bases with political jargon and appealing platforms. This was seen in the Democratic governor's push for SOL reform in 2014 and the Republican governor's push for excellence in 2023. The rhetoric of these campaign promises have *not* matched the putative outcomes. They have, however, resulted in political capital for the politicians, meaning that while the students should have been the winners of the policy, really it was the politicians who won.

These politics result in the production of unclear policies. While policy is often considered a technical and rational process, the results of this CPA show how infused with values and politics the policy process is. It is not only the ends of policy nor solely the means of

policy, but the whole policy process that is infused with politics and values. The political back and forth within the policy formulation leads to dysfunction for implementers and for further policy decisions by policy makers as the substantive aspects of the policy are kicked down the road.

There is a theoretical stance that policies create politics (e.g., Lowi, 1972); however this CPA aligns with researchers (e.g., Ingram & Schneider, 2006; Smith et al., 2004) who assert the opposite, that politics are infused into policies. As Smith et al. (2004) explain:

An idealized relationship between politics and policy assumes that mutual, good-faith persuasion and tradeoffs between groups having roughly equivalent power will yield good policy outcomes that are fair to both interest groups and have reasonable and obvious expectations for certain benefits and costs— both to those groups and to society as a whole. Now, how does this version of policy stand up to what we know about contemporary American politics? Not very well. (p.4)

Ingram and Schnedier (2006) have coined the actual policy process, in contrast to the ideal one described in the quote above, as *degenerative politics*; they explain that "Problems do not just happen. They are constructed through the interaction of a variety of political phenomena including existing public policies" (Ingram & Schneider, 2006, p.174). So, policymakers are not only in the business of solving problems, but also creating them. This has been true of the entire system of standards, assessment, and accountability policy at the national and state levels from the 1980s forward and in Virginia in this CPA.

In standards, the creation of a problem occurred by pointing to typographical errors in the History & Social Studies standards that were to be fixed, and then rather than addressing those minor edits, introducing an entirely new set of standards at a following meeting of the Board. In

assessment, problem creation was centered on the discussion of SOL cut scores as well as creating concern around the "lowering the standards" without evidence to how this impacted student achievement. Lastly, problem construction was seen in the Virginia BOE's discussion around accountability, when they were encouraged to *create* urgency by changing how schools were labeled in the accountability policy to an A-F grade. Each of these instances helps illuminate that the policy process is not totally rational but is infused with varying perspectives often fueled by politics external to education policy. Here is where the concept of *political spectacle* (Edelman, 1988, as cited in Smith et al., 2004) is helpful to understand Virginia's standards, assessment, and accountability policies.

Education policy, and thus education itself, then becomes a vessel for political gain, made through theatrical and symbolic displays that really have little to do with education. This occurs when the discussion around policy decisions is not rooted in the actual outcomes nor the well-being of students. Thus, political gain thus becomes an additional purpose of public education from a state politician's perspective, along with the myriad of other purposes assigned to public schools by the policy process (as described in Table 11 above). This is supported by Smith et al. (2004) who claim that "American politics have become detached from their democratic foundations and how these conditions of politics distort public policies, especially education policies" (p.2). Examples from this research support those provided by Smith and colleagues (2004), providing evidence that policies born out of these degenerative politics "create perverse consequences in schools, frustration and perplexity among the public" (p.5) leading to additional mixed messaging at all levels of the policy process.

Challenging the Status Quo

Rein (1976) encourages the policy analyst to "distrust orthodoxy," and to be on vigilant lookout for those spaces within policy and practice where alternative policy solutions to those currently in use would better meet the needs of the intended audience (p. 25). The above description of degenerative politics indicates that Virginia's standards, assessment, and accountability system is one that should inspire a healthy skepticism. Young and Diem (2017) advise looking for where those who are oppressed by the system are resisting it. Often resisting practitioners and advocates are a knowledgeable guide to alternative pathways. The system itself is resistant to change, as seen by the continuing dominance of rote standards and standardized testing for accountability purposes driven by Virginia policymakers. However, teachers, business leaders, and students are showing up at Board meetings and public comment and writing to local media outlets to advocate for the 21st century skills aligned with the Profile of a Graduate. As discussed in papers two and three, teachers see the value in these skills and aspire to be part of a system that prioritizes them.

If the BOE wants to take seriously their charge to supervise a public system of education in a way that meets their self-identified number one priority to "provide high-quality, effective learning environments for all students" (VDOE, 2023b), then considerations about how the system of standards, assessment, and accountability as-is fail to do this is warranted. As next steps, these policymakers must consider how to align the new standards with assessments and accountability and how to better create clearer messaging that would create effective learning environments for all students. Policymakers should consider: What would a system of standards, accountability, and assessment look like if the policymakers were committed first to the Profile and the 5 C's and the appropriate manner of performance assessments to measure them? How

would the accountability system need to be reconfigured to prioritize those standards and assessment? What messages would need to be sent by policy to educators who must implement those new priorities? In order to make change in a system deeply rooted in standardized thinking, policymakers will need to confront the ways that politics are a barrier to addressing the needs of students, and messaging will need to be more purposeful and clearer for implementers.

Role as a Moral Critic - Future Recommendations

In my role as a policy analyst and moral critic (Rein, 1976), I questioned if the decisions that led to the current system were really made with the priority, to "provide high-quality, effective learning environments for all students" (VDOE, 2023b). Based on this research I would urge the re-centering of this value in future BOE decision-making processes. As individuals, Board members and VDOE staff have the opportunity to reframe their political perspectives in terms that are humanizing for Virginia students. Innovative approaches to accountability, such as sampling students or testing every few years, would be less expensive and allow for less time focused on testing and more time that focused on the authentic learning of 21st century skills. Conversations around *creating* urgency provide evidence that the high stakes accountability system prioritizes symbolic competition against other states over authentic educational quality or equity outcomes for students. Other states have managed to shift to higher order standards and performance assessment and others are experimenting with this approach. Virginia, on the other hand, has been incapable of making this transition because it has yet to move away from the foundational belief in the need to hold everyone accountable to uniform "objective" measurement. For further reflections on Flyvbjerg's (2001) questions "What should we do?" (pp. 136-7), recommendations, and next steps for action, see this dissertation's epilogue.

Future Implications & Further Research

Unfortunately, the findings of this CPA have largely echoed the findings from a study in the 2010s that reported that Virginia school systems are *not* ensuring high-quality, effective learning environments for all students. In investigating student achievement outcomes utilizing data from districts in Virginia in 2010, authors report:

These disparate educational experiences and outcomes strongly suggest that providing equality of educational opportunity for all children remains a compelling social problem that has yet to be adequately solved and are exasperated by accountability systems based on standardized testing that discount the reality of segregated and resource-depleted school environments that face black students in Virginia disproportionally more than their white peers. Until these systems are changed and commitment to a truly equal public school experience is fulfilled, racial disparities will only grow more apparent in Virginia's as well as the nation's schools. (Brunn-Bevel & Byrd, 2015, p.444)

While over ten years have passed, even with increased rhetoric about alternative standards, assessment, and accountability approaches, the Virginia system has muted the possibilities of how new initiatives, such as performance assessments associated with the Profile, could impact classroom teaching. As papers two and three investigate, when such ambiguous policies are interpreted and translated at the district and teacher level, the great influence of the standardized assessment system prevails and causes dissonance with teachers who are trying to deliver humane educational practice. While teachers' own philosophies often lean more toward constructivist approaches to standards, instruction, and assessment that allow for a growth mindset, the system itself holds their feet to the fire with accountability structures that require

rote teaching and drill and kill pedagogies. Even the aspirations of some forward-looking policymakers are dashed within the prevailing values of efficiency.

Further research is needed to understand how state-level policy actors respond to these findings. When faced with the political spectacle, are they aware of how much it impacts them, or are they, too, so steeped in the system that they are unaware of the ways in which they are pawns at play? While observation of interactions between BOE and VDOE members illuminated some perspectives, the lack of interview data was a limitation of this study. Further research that highlights legislator, BOE, and VDOE staff perspectives could help determine how policy actors see their roles in the system and lead to further discoveries of places for innovation and resistance. Additionally, research in other contexts that have shifted from standards-based to performance-based systems successfully would allow for interesting perspectives about policy transformation. Research that centered student perspectives in those systems could help illuminate system impacts that may not be clear in standardized measurements often used to measure student achievement and success.

Significance and Conclusions

This CPA illuminates the negative role of politics in the role of public schools. While politicians will claim that the children are our future, personal victories motivate much of the educational agenda at the state level. This CPA encourages all citizens to take a critical perspective when listening to the policy promises of any politician and reveals the need to look not only at policy rhetoric but also policy outcomes. Policy cannot be taken at face-value, and policy outcomes that are divergent from policy goals are not only mis-translations by implementers, but often intentional virtue signaling without substance.

The results of this CPA may inspire in the reader a feeling of pessimism in the policy process; however, the notions of resistance can inspire hope that there may be a brighter future. Students and teachers who are engaging in the policy process and pushing back on the politicization of their educations and careers point to a possibility of naming racist and classist policies and addressing issues related to race and class head-on rather than from a saviorism mentality. For next steps and suggestions for this process, see the epilogue.

As explored in this CPA, some members of the Board seem eager to think about new systems, and are aware of the ways that standards, assessment, and accountability are linked. There are legislators, members of the VDOE, and Board members who are open to exploring new systems. As these changes are considered, the political spectacle, especially around politically sensitive issues such as race, will need to be addressed and overcome to allow for clear messaging. This clear messaging will be imperative if the new system is to have the impact on all Virginia students that all leadership claims it hopes to see.

References

Alderson, S. (2022, May 22). Spencer Alderson column: An alternative to Standards of Learning tests. *Richmond Times-Dispatch*. https://richmond.com/opinion/columnists/spencer-alderson-column-an-alternative-to-standards-of-learning-tests/article_054900db-fefd-5327-ba90-a11a0633e9b8.html

Alfuth, J. (2021). Assessments of and for learning: How states are rethinking accountability and assessment policies in the COVID-19 era. *KnowledgeWorks*. https://knowledgeworks.org/resources/assessments-accountability-learning-states-covid-19/

Altheide, D. L., & Schneider, C. J. (2013). *Qualitative media analysis*. SAGE.

Atkinson, D. (2017). Virginia rethinks high school in its Profile of a Graduate. *State Education Standard*, *17*(2), 28–33. National Association of State Boards of Education.

Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy*, *30*(1), 39–62.

Ayers, E. (2023, March 9). Ed Ayers: Let educators, not politicians, shape K-12 education. *Richmond Times-Dispatch.* https://richmond.com/opinion/columnists/ed-ayers-let-educators-not-politicians-shape-k-12-education/article_a72dd972-bdcd-11ed-b448-af3862baf1c0.html

Bass, S. (2022, October 30). Column: Politicizing test scores won't solve the education crisis. *Richmond Times-Dispatch https://richmond.com/opinion/columnists/column-politicizing-test-scores-won-t-solve-the-education-crisis/article_48222f74-1081-5cf6-a551-77df21949711.html*

Beadie, N. (2004). Moral errors and strategic mistakes: Lessons from the history of student accountability. *Holding accountability accountable: What ought to matter in public education*, 35-50.

Bennett, J. (2019). Removing race: How context and colorblindness influence conceptualizations of equity in a third-grade rural classroom. *Race and Pedagogy Journal, 4*(1).

Berliner, D. C., & Biddle, B. J. (1997). *The manufactured crisis: Myths, fraud, and the attack on America's public schools* (collegiate edition). Addison-Wesley.

Bradbury, A. (2020). A critical race theory framework for education policy analysis: The case of bilingual learners and assessment policy in England. *Race, Ethnicity and Education*, *23*(2), 241-260.

Brunn-Bevel, R. J., & Byrd, W. C. (2015). The foundation of racial disparities in the standardized testing era: The impact of school segregation and the assault on public education in Virginia. *Humanity & Society*, *39*(4), 419-448.

Bryson, A. (2022a, October 28). Virginia's Black and Latino students hit hardest by pandemic learning loss. *Richmond Times-Dispatch*. https://richmond.com/news/local/education/virginia-s-black-and-latino-students-hit-hardest-by-pandemic-learning-loss/article_e7334916-7cf1-56ee-b343-04d837bca9b3.html

Bryson, A. (2022b, November 16). New draft history standards reorient framing of race relations. *Richmond Times-Dispatch.*

Bryson, A. (2022c, November 17). Board of Education rejects Youngkin's proposed revisions to K-12 history standards. *Richmond Times-Dispatch.*

Cannell, J. J. (1989). How public educators cheat on standardized achievement tests: The "Lake

    Wobegon" report.

Chandler, M.A. (2014a, April 5). Virginia students will take fewer Standards of Learning tests

    next year. *Washington Post*. https://www.washingtonpost.com/local/education/virginia-

    students-will-take-fewer-standards-of-learning-tests-next-year/2014/04/05/eea18666-

    bb46-11e3-9a05-c739f29ccb08_story.html

Chandler, M.A. (2014b, April 5). What are 'authentic assessments'? How are they different from

    multiple-choice tests? *Washington Post*.

    https://www.washingtonpost.com/local/education/virginia-students-will-take-fewer-

    standards-of-learning-tests-next-year/2014/04/05/eea18666-bb46-11e3-9a05-

    c739f29ccb08_story.html

Chen, G. (2021, May 13). "When teachers cheat: The standardized test controversies" *Public

    School Review. https://www.publicschoolreview.com/blog/when-teachers-cheat-the-

    standardized-test-controversies*

Cohen, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform works*. New

    Haven, CT: Yale University Press.

Corbishley, J. B., & Truxaw, M. P. (2010). Mathematical readiness of entering college freshmen:

    An exploration of perceptions of mathematics faculty. *School Science and Mathematics*,

    110(2), 71-85.

Coyner, C. (2021, January 10). Carrie Coyner column: Covid-19 shows that student testing must

    evolve: It's time for growth based assessments in Virginia. *Richmond Times-Dispatch.*

    https://richmond.com/opinion/columnists/carrie-coyner-column-covid-19-shows-that-

student-testing-must-evolve-its-time-for-growth/article_9de42480-0b2e-55bd-8c23-2014536682c8.html

Culpepper-Ofori, D. (2012). *The development of tracking and its historical impact on minority students (Publication No.*3509468) [Doctoral dissertation, Walden University]. ProQuest Dissertations Publishing.

Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives, 22*(86),1.

Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and management*, *30*(3), 418-446.

Duke, D. L., & Reck, B. L. (2003). The evolution of educational accountability in the Old Dominion. In D.L Duke, M. Grogan, P. Tucker, & W.F. Heinecke, (Eds.), E*ducational leadership in an age of accountability: The Virginia experience*, (pp. 36-68). SUNY Press.

Elementary and Secondary Education Act of 1965: H. R. 2362, 89th Cong., 1st sess., Public law 89-10. Reports, bills, debate and act. (1965).

Er, S. N. (2018). Mathematics readiness of first-year college students and missing necessary skills: perspectives of mathematics faculty. *Journal of Further and Higher Education*, *42*(7), 937-952.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119–161). New York, NY: Macmillan.

Evans, C. M. (2019). Effects of New Hampshire's innovative assessment and accountability system on student achievement outcomes after three years. *Education Policy Analysis Archives*, *27*(10), 1-37.

Evans, K.B., & Remmers, V. (2017, July 19). Several Petersburg school employees fired

    following investigation into SOL cheating. *Richmond-Times Dispatch.*

    https://richmond.com/news/local/central-virginia/several-petersburg-school-employees-

    fired-following-investigation-into-sol-cheating/article_dcbcb051-f843-59cc-ae31-

    11c88f1575b3.html

Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).

Ferlazzo, L. (2021, Feb 02). Students respond to adults' fixation on 'learning loss.'

    EducationWeek. https://www.edweek.org/leadership/opinion-students-respond-to-adults-

    fixation-on-learning-loss/2021/02

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can*

    succeed again. Cambridge University Press.

Gardner, H. (1983). *The theory of multiple intelligences*. Heinemann.

Goldin, S., & Khasnabis, D. (2021). In the pursuit of justice: Moving past color-evasive efforts.

    The Educational Forum, 86(1), 1-4.

Gould, S. J. (1996). *The mismeasure of man: Revised and expanded edition.* W. W. Norton &

    Company.

Grissmer, D., & Flanagan, A. (1998). Exploring rapid achievement gains in North Carolina and

    Texas. *Lessons from the States*. National Education Goals Panel.

Guba, E. G. (1984). The effect of definitions of "policy" on the nature and outcomes of policy

    analysis. *Educational Leadership*, *42*(2), 63-70.

Heinecke, W. F., Curry-Corcoran, D. E., & Moon, T. R. (2003). US schools and the new

    standards and accountability initiative. *Educational leadership in an age of*

    accountability: The Virginia experience, 7-35.

Hollingworth, L. (2009). Unintended Educational and Social Consequences of the No Child Left

    Behind Act. *The Journal of Gender, Race, and Justice*, *12*(2), 311.

HB 585, Chapter 760 (2022) (enacted). https://lis.virginia.gov/cgi-

    bin/legp604.exe?221+ful+CHAP0760

Hunter, K. (2021, September 14). Richmond students are behind, according to SOL data, but low

    testing participation makes the data hard to judge. *Richmond Times-Dispatch.*

    https://richmond.com/news/local/education/richmond-students-are-behind-according-to-

    sol-data-but-low-testing-participation-makes-the-data/article_1400749f-0807-5dd1-

    a8b5-54202b15d5ab.html

Hyde, A. M. (2007). Self-constitution as resistance to normalization: Educator agency in the era

    of accountability (Doctoral dissertation, University of Pittsburgh).

IDRA (2022). What Virginia's anti-equity Executive Order 1 and reports mean for K-12 schools

    and students: A guide for school leaders. https://www.idra.org/wp-

    content/uploads/2022/11/What-Virginias-Anti-Equity-EO1-Means-IDRA-2022-Nov.pdf

Ingram, H., & Schneider, A. L. (2006) Policy analysis for democracy. In M. Moran, M. Rein &

    R.E. Goodin (Eds.), *The Oxford handbook of public policy* (pp.169-189). Oxford

    University Press.

Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for*

    a socially just future. Parlor Press LLC.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing

    in the Chicago Public Schools. *Journal of public Economics*, *89*(5-6), 761-796.

Janesick, V. J. (2016). *Stretching exercises for qualitative researchers.* Sage.

Knecht, D. (2007). The consortium and the commissioner: A grass roots tale of fighting high stakes graduation testing in New York. *The Urban Review*, *39*(1), 45–65.

Koretz, D. M., Barron, S., Mitchell, K.J., & Stecher, B.M. (1996). *Perceived effects of the Kentucky Instructional Results Information System (KIRIS)*. Santa Monica, CA: Rand.

Koretz, D. (1998). Large-scale portfolio assessments in the US: Evidence pertaining to the quality of measurement. *Assessment in Education: Principles, Policy & Practice*, *5*(3), 309–334.

Labaree, D. F. (1997). Public goods, private goods: The American struggle over educational goals. *American Educational Research Journal*, *34*(1), 39-81.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational Researcher*, *35*(7), 3-12.

Leonhardt, D. (2022, May 2). Why did U.S. schools make so much progress in the 1990s and early 2000s? *The New York Times: The Morning*. https://www.nytimes.com/2022/05/02/briefing/school-education-gaps-funding-math.html

Leonor, M. (2022, March 10). School superintendents chide Youngkin administration's audit of 'divisiveness' in schools. *Richmond Times-Dispatch*. https://richmond.com/news/state-and-regional/govt-and-politics/school-superintendents-chide-youngkin-administrations-audit-of-divisiveness-in-schools/article_80569e29-95cf-5ab7-95a1-bb2becd72640.html

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2018). Paradigmatic controversies, contradictions, and emerging confluences revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (5th ed., pp. 108-150). Sage.

Lipsitz, G. (2006). *The possessive investment in whiteness: How white people profit from identity politics*. Temple University Press.

Llovio, L. Governor signs legislation to transform Va. high school graduation requirements.

    Richmond Times-Dispatch. https://richmond.com/news/local/governor-signs-legislation-

    to-transform-va-high-school-graduation-requirements/article_ff9ba6c9-14ed-5745-a8ac-

    898e8744991a.html

Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of*

    educational freedom. Beacon Press.

Lowi, T. J. (1972). Four systems of policy, politics, and choice. *Public Administration Review,*

    32(4), 298-310.

Mabry, L. (2004). Strange, yet familiar: Assessment driven education. *Holding accountability*

    accountable: What ought to matter in public education, 116-134.

Mahoney, J. (2020, May 7). Editorial: Measuring student progress in a post-pandemic world.

    Richmond Times-Dispatch. https://richmond.com/opinion/editorial/editorial-measuring-

    student-progress-in-a-post-pandemic-world/article_88bc3cc7-54b6-5a1a-8d0e-

    d771d48bca88.html

Marchant, G. J., Paulson, S. E., & Shunk, A. (2006). Relationships between high-stakes testing

    policies and student achievement after controlling for demographic factors in aggregated

    data. *Education Policy Analysis Archives*, *14*(30), 1–31.

Marshall, C., Mitchell, D., & Wirt, F. (1989). *Culture and education policy in the American*

    states. The Falmer Press.

Mattingly, M. (2018, August 17). Virginia Department of Education to release SOL test results

    Wednesday amid questions about the data's value; 'We haven't fixed things, we've made a

    bigger mess'. *Richmond Times-Dispatch.* https://richmond.com/news/local/virginia-

department-of-education-to-release-sol-test-results-wednesday-amid-questions-about-the-datas/article_d665548d-0437-5bc7-aa51-3d83817f804b.html

McCann, M. (2021, April 20). *Response to information request*. Education Commission of the United States. https://www.ecs.org/wp-content/uploads/State-Information-Request-_-2021-Assessment-and-Accountability-Waivers.pdf

McDonnell, L. M. (2005). Assessment and accountability from the policymaker's perspective. In J. L. Herman & E. H. Haertel (Eds.), *Uses and misuses of data for educational accountability and improvement. The 104th yearbook of the National Society for the Study of Education, Part II* (pp. 35-54). Blackwell.

Mervosh, S. (2022, Nov 28). *Pandemic learning loss. New York Times: The Morning.* https://www.nytimes.com/2022/11/28/briefing/pandemic-learning-loss.html

Mirshahi, D. & DeFusco, J. (2023, January 6). Virginia Department of Education unveils new K-12 history standards draft. WRIC. https://www.wric.com/news/virginia-news/virginia-education-department-unveils-new-k-12-history-standards-draft/

Moore, M. (2023, March 10). Virginians respond to Youngkin's CNN Education Town Hall with fact-checking, skepticism. *Dogwood*. https://vadogwood.com/2023/03/10/virginians-respond-to-youngkins-cnn-education-town-hall-with-fact-checking-skepticism/

Natanson, H. & Vozzella, L. (2022, May 19). Virginia Dept. of Education releases report on student achievement. *Washington Post.* *https://www.washingtonpost.com/education/2022/05/19/virginia-schools-student-achievement-youngkin/*

National Association of Colleges and Employers (2019). *NACE Job Outlook 2020*. www.naceweb.org

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for

educational reform. (ED 226006). ERIC. https://files.eric.ed.gov/fulltext/ED226006.pdf

Newmann, F. M., King, M. B., & Rigdon, M. (1997). Accountability and school performance:

Implications from restructuring schools. *Harvard Educational Review*, 67, 41–74.

Nichols, S. L. (2007). High-stakes testing: Does it increase achievement? *Journal of Applied

School Psychology*, *23*(2), 47–64.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts

America's schools*. Harvard Education Press.

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

https://www.govinfo.gov/content/pkg/PLAW-107publ110/pdf/PLAW-107publ110.pdf

Nocera, J. (2022, August 18). Virginia students' performance on SOLS still trails pre-pandemic

scores. *Richmond-Times Dispatch*. https://richmond.com/news/local/education/virginia-

students-performance-on-sols-still-trails-pre-pandemic-scores/article_d8bada18-4eee-

578c-8efc-32bad8b10506.html

Nomberg, R. (2022, June 18). Rob Nomberg column: Prioritizing the development of durable

skills. *Richmond Times-Dispatch. https://richmond.com/opinion/columnists/rob-

nomberg-column-prioritizing-the-development-of-durable-skills/article_9bcb2648-a2ac-

5bb4-890d-7fe8debe9928.html*

Prasse, D. P., & Reschly, D. J. (1986). Larry P.: A case of segregation, testing, or program

efficacy?. *Exceptional Children*, *52*(4), 333-346.

Prunty, J. J. (1985). Signposts for a critical educational policy analysis. *Australian Journal of

Education*, *29*(2), 133-140.

Randall, J. (2021). "Color-neutral" is not a thing: Redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educational Measurement: Issues and Practice*, *40*(4), 82-90.

Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting White supremacy in assessment: Toward a justice-oriented, antiracist validity framework. *Educational Assessment*, 1-9.

Rein, M. (1976). *Social science and public policy.* Penguin Books.

Rothman, R., & Marion, S. F. (2016). The next generation of state assessment and accountability. *Phi Delta Kappan*, *97*(8), 34-37.

Ruff, R. R. (2019). State-level autonomy in the era of accountability: A comparative analysis of Virginia and Nebraska education policy through No Child Left Behind. *Education Policy Analysis Archives* (5/6), 1–28.

Rusk, D. (1998). To improve poor children's test scores, move poor families. *Abell Report* Baltimore, MD: Abell Foundation.

Sabiston, N. (2016). What an SOL test opt out means for you and your school. *Virginia is for Teachers*. https://www.virginiaisforteachers.com/2016/06/what-sol-test-opt-out-means-for-you-and.html

Schwartz, R. B., Robinson, M. A., Kirst, M. W., & Kirp, D. L. (2000). Goals 2000 and the standards movement. *Brookings Papers on Education Policy,* (3), 173-214.

Smith, M. L. (1991). Meanings of test preparation. *American Educational Research Journal*, *28*(3), 521-542.

Smith, M. L., Edelsky, C., Draper, K., Rottenburg, C., & Cherland, M. (1990). The role of testing in elementary schools (CSE Tech. Rep. No. 321). Los Angeles: University of California. *Center for Research on Evaluation, Standards, and Student Testing*.

Smith, M. L., & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education*, *51*(5), 334-344.

Smith, M. L., Miller-Kahn, L, Heinecke, W. & Jarvis, P.F. (2004). *Political spectacle and the fate of American schools*. New York: Routledge.

Smith, M. L., Noble, A. J., Cabay, M., Heinecke, W., Junker, M. S., & Saffron, Y. (1994). What happens when the test mandate changes? Results of a multiple case study. *CRESST Report*, *380*.

Spillane, J. P. (2005). *Standards deviation*. Harvard University Press.

Standards of Accreditation, 8 Virginia Administrative Code. §20-131 (2018). https://law.lis.virginia.gov/admincode/title8/agency20/chapter131/

Standards of Quality, Code of Virginia. §22.1-253.13.2 (2014).

Standards of Quality, Code of Virginia. §22.1-253.13.2 (2018).

Standards of Quality, Code of Virginia. §22.1-253.13.2 (2021). https://law.lis.virginia.gov/vacode/title22.1/chapter13.2/

Stecher, B. (2010). *Performance assessment in an era of standards-based educational accountability*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.

Stein, S. J. (2004). *The culture of education policy*. Teachers College Press.

Stewart, C., Wall, A., & Marciniec, S. (2016). Mixed signals: Do college graduates have the soft skills that employers want?. *Competition Forum, 14(2)*, 276. American Society for Competitiveness.

Stosich, E. L., Snyder, J., & Wilczak, K. (2018). How do states integrate performance assessment in their systems of assessment? *Education Policy Analysis Archives*, *26*, 13.

Tracey, C. (2005). Listening to teachers: Classroom realities and NCLB. In G. L. Sunderman, J.S. Kim, & G. Orfield (Eds.), *NCLB meets school realities: Lessons from the field (pp. 81-103)*. Corwin Press.

Turner, C. (2022, June 22). 6 things we've learned about how the pandemic disrupted learning. *NPR: All Things Considered.* https://www.npr.org/2022/06/22/1105970186/pandemic-learning-loss-findings

Virginia Board of Education. (2019, January 24). Guidelines for local alternative assessments for 2018-2019 through 2019-2020. https://www.doe.virginia.gov/home/showpublisheddocument/20558/638043628349270000

Virginia Board of Education. (2021, June 17). Guidelines for local alternative assessments: 2021-2022 and beyond. https://www.doe.virginia.gov/home/showpublisheddocument/42537/638113716617030000

Virginia Board of Education. (2023). Proposed revised 2023 History and Social Science Standards of Learning. Virginia Department of Education. https://www.doe.virginia.gov/home/showpublisheddocument/42385/638103461592970000

Virginia Department of Education (n.d.). Implementation support for balanced assessment plans.

https://www.doe.virginia.gov/home/showpublisheddocument/42515/638113600284870000

Virginia Department of Education (2017a, September 26). Revised Standards of Accreditation:

The case for reform [Video]. YouTube.

https://www.youtube.com/watch?v=uUqQZQApmUE&list=PLRTyI0-

OTuVOaWG7r0o2WLG7xdes6pS5H&index=2

Virginia Department of Education (2017b, September 26). Revised Standards of Accreditation:

Continuous improvement for all schools. YouTube.

https://www.youtube.com/watch?v=_ksyK_BzUk8&list=PLRTyI0-

OTuVOaWG7r0o2WLG7xdes6pS5H&index=4

Virginia Department of Education (2017c, September 26). Revised Standards of Accreditation:

Introduction and overview [Video]. YouTube.

https://www.youtube.com/watch?v=1WSTZ111VKM&list=PLRTyI0-

OTuVOaWG7r0o2WLG7xdes6pS5H&index=1

Virginia Department of Education (2018a, May 30). Performance assessments [Video].

YouTube. https://www.youtube.com/watch?v=aDV7CywKLYo&list=PLRTyI0-

OTuVNK2ataXB_JQOHzCgRjuYZq&index=2

Virginia Department of Education (2018b, May 30). Regional performance assessment training

[Video]. YouTube. https://www.youtube.com/watch?v=APFnaB4OH2Q&list=PLRTyI0-

OTuVNK2ataXB_JQOHzCgRjuYZq&index=1

Virginia Department of Education (2022). Profile of a Virginia Graduate. Virginia Department of

Education. https://www.doe.virginia.gov/parents-students/for-students/graduation/policy-

initiatives/profile-of-a-virginia-graduate

Virginia Department of Education (2023a). Participation & inclusion. Virginia Department of

    Education. https://www.doe.virginia.gov/teaching-learning-assessment/student-

    assessment/virginia-sol-assessment-program/participation-inclusion

Virginia Department of Education (2023b). Virginia Board of Education. Virginia Department of

    Education. https://www.doe.virginia.gov/data-policy-funding/virginia-board-of-

    education

Virginia Department of Education (2023c). Virginia Department of Education Code of Ethics.

    Virginia Department of Education. https://www.doe.virginia.gov/about-vdoe/virginia-

    department-of-education-code-of-ethics

Vozzella, L. & Asbury, N. (2022, September 23). Youngkin wants a better measuring stick for

    Va. schools. What's his plan? *Washington Post.* https://www.washingtonpost.com/dc-

    md-va/2022/09/23/youngkin-school-ratings/

Wasserberg, M. J. (2017). High-achieving African American elementary students' perspectives

    on standardized testing and stereotypes. *Journal of Negro Education*, *86*(1), 40-51.

Weiner, K. G., & Carter, P. L. (2013). Achievement gaps arise from opportunity gaps. *Closing*

    the opportunity gap: What America must do to give every child an even chance, 1(10).

Williams, M. (2021, May 21). More parents opting students out of SOL tests. *Daily News*

    Record. https://www.dnronline.com/news/rockingham_county/more-parents-opting-

    students-out-of-sol-tests/article_d3d19447-e159-5e12-bf3a-eff23a6b5c81.html

Wilt, C. L., Annamma, S. A., Wilmot, J. M., Nyegenye, S. N., Miller, A. L., & Jackson, E. E.

    (2022). Performing color-evasiveness: A DisCrit analysis of educators' discourse in the

    US. *Teaching and Teacher Education*, *117*, 103761.

Yilmaz, M., Buchanan, M. B., & Martin, J. (2023). Disrupting majoritarian narratives in high stakes testing: A critical race theory analysis of assessment. *Equity in Education & Society*, *2*(1), 28-42.

Young, M. D. (1999). Multifocal educational policy research: Toward a method for enhancing traditional educational policy studies. *American Educational Research Journal*, *36*(4), 677-714.

Young, M. D., & Diem, S. (2017). Introduction: Critical approaches to education policy analysis. In M.D. Young & S. Diem (Eds). *Critical approaches to education policy analysis: Moving beyond tradition*, 1-13.

Young, M. D., & Reynolds, A. L. (2017). Critically examining policy workers and policy work within state boards of education. *Critical approaches to education policy analysis: Moving beyond tradition*, 19-41.

Youngkin, G. (2023, March 9). *Transcript from CNN Town Hall 'The War Over Education.'* (J. Tapper, Interviewer). Retrieved from https://transcripts.cnn.com/show/se/date/2023-03-09/segment/01

Paper Two: The Standards of Learning and the 5 C's: Virginia State Policies in Action

In education there are frequently policy directives from federal and state Departments of Education (DOE) that dictate the content that teachers must include in their instruction of students. In Virginia, two of these policies include the Standards of Learning, assessed by standardized, high-stakes tests, and the Profile of a Virginia Graduate, which includes a mandate requiring that students graduate with the 5 C's: collaboration, communication, citizenship, creative thinking, and critical thinking skills. All governmental policies are based in societal values, but sometimes those values contradict each other. When this happens, the conflict is mitigated in how the policies are implemented. Guba (1984) asserted that policy-in-intention often differs from policy-in-action. In the federal system of the United States, policies are being interpreted at every level of the system, which often changes the ways that policies get implemented. The purpose of this research was to analyze, from the perspectives of various stakeholders in the Virginia public education system, how these two state policies are being interpreted and implemented.

This research has the potential to help policymakers and practitioners learn about how the two potentially conflicting policies are impacting education in the state, changes that may need to be made, and improvements of the policy system.

As the focus of the research was on the meaning-making, the research questions and design for this project were shaped by the interpretive paradigm with the understanding that individuals have diverse experiences and ways of understanding what is happening around them. An interpretive qualitative case study design was employed because it is useful in understanding how stakeholders at different levels and positions within the education system understand and

utilize the two policies being considered: The Standards of Learning and the 5 C's. Data was analyzed through an interpretive frame of analytic induction (Erickson, 1986).

Background

The purpose of this study was to learn more about the experiences of English teachers with regards to the Virginia Standards of Learning (SOLs) and the 5 C's in order to understand how these policies are being translated and implemented. The Standards of Learning are utilized both as a means for structuring the curriculum within Virginia and for assessment and accountability purposes. Each subject and grade has a curriculum framework provided by the state that explains what should be covered under the standards, and certain courses throughout elementary, middle, and high school are assessed with SOL tests that are used to hold students, teachers, and schools accountable. In 2016, the Virginia legislature mandated that schools implement the Profile of a Virginia Graduate alongside the Standards of Learning, which focused more on higher order thinking skills. Included in this policy mandate were the 5 C's focusing on citizenship, communication, collaboration, creative, and critical thinking skills. In this study, I investigate, within the Virginia public school system, stakeholders' meaning-making around policies for standardized testing and soft skills acquisition through a qualitative case study design. Data collection for this project began right at the beginning of stay-at-home orders in the state of Virginia due to the COVID-19 pandemic. This required an adaptation of research methods but provided an interesting environment for data collection as the end-of-course tests associated with the Standards of Learning were ultimately canceled during the 2019-2020 academic year.

Brief Review of Relevant Literature

When researchers, parents, and policy makers consider the success of American public schools, they usually prioritize two indicators in evaluating achievement: math and reading test scores. Most people agree, however, that schools are responsible for teaching students more than how to pass math and reading tests. In the early 1990s, North Carolina and Texas led the way in the implementation of high-stakes testing program (Grissmer & Flanagan, 1998). Since that time, a national regime of high stakes accountability has been the largest influence in education policy nationally. Following this national movement, Virginia implemented the Standards of Learning (SOL) as their form of high-stakes accountability. Researchers have been challenging the connection between high-stakes testing and school performance since before the implementation of No Child Left Behind (NCLB) in the United States (Newmann et al., 1997). After NCLB was enacted in 2002, additional groups of researchers analyzed data comparing scores in states with high-stakes and low-stakes testing across the U.S. (Braun, 2004; Nichols, Glass, & Berliner, 2006, 2012). Depending on their methods, authors often reported that high-stakes accountability practices had varying levels of impact on scores, but few analyses demonstrated that high-stakes testing functioned well as an accountability system. Although reading and math are essential subjects to attain proficiency in, there are numerous other skills that students must attain to be prepared for society.

Recent data also supports that high-stakes testing is no longer associated with the same levels of growth in educational outcomes as asserted in the early days of the movement. In 2019, aside from a one-point gain in fourth-grade math scores, students' reading and math scores on NAEP declined on average nation-wide (Nation's Report Card, 2019). In 2017, Virginia implemented the Profile of a Virginia Graduate to encourage the development of skills other than

those assessed on standardized testing. The motivation behind this Profile was to bridge students'

experiences in schools with what they would experience after graduation, either in their career or

post-secondary education. During the development of the Profile, an interviewed member of the

Department of Education shared that community partners such as stakeholders from higher

education, the business community, and industry partners were met with in order to make sure

the necessary skills were included. The Profile, the Board of Education states, "describes the

knowledge, skills, experiences and attributes that students must attain to be successful in college

and/or the work force and to be 'life ready'" (Virginia Department of Education, 2020). This

profile includes the policy directive that students should learn the 5 C's: critical thinking,

creative thinking, collaboration, communication, and citizenship.

While both the Standards of Learning and the 5 C's policy frameworks were mandated at

the state level in Virginia, there is less available data about how the policies are being understood

and interpreted by teachers within the system. Due to the fact that the policies focus on different

sets of skills and are monitored in different ways, with only one tied to high-stakes testing, there

is potential conflict between the emphasis and priority-signaling associated with these policies

that could impact their implementation. This study gathered information from teachers and

employees of the Virginia Department of Education in order to help illuminate the differences

between policy-in-intention and policy-in-action as highlighted and defined by Guba (1984). As

Hall and McGinty (1997) phrase it, "Policies are vehicles for the realization of intentions" (p.

441.) However, the intentions and interpretations of all the stakeholders who interact with a

policy also influence what happens because of that policy. It is easy to understand this in the

context of education, where Boards of Education and Virginia legislators are far removed from

teachers who are interacting with students and enacting the policies throughout the school year.

Research Questions

The research questions addressed in this study were:

1. How do stakeholders make meaning of varied state policies?

2. How do teachers decide how to implement state policies?

3. How are policy intentions similar to and different from policy outcomes?

Methodology

In order to address the research questions centered on policy meaning, this study was conducted within the interpretive paradigm. Due to the centrality of meaning-making to this project, this paradigm, which prioritizes research participants' understanding of reality, was best aligned with the research questions being asked. An interpretive case study using the research strategy of analytic induction (Erickson, 1986) was employed.

Methods

Virginia was chosen as a site of research due to the potential conflict between the two state policies in implementation. For the micro-level of this research project, one rural school division that had moderate diversity in race and socio-economic status across the district was selected. The demographics of each school are presented in Figure 1.

Due to the SOL testing requirements for English courses in K-12, and the likelihood that the 5 C's would also be a focus within instruction, English courses were used to investigate these differences in implementation. An elementary, middle, and high school located within the same town within the division were chosen by convenience sampling at the time the study was designed. The charts along the right side of Figure 1 below indicate how students in each of the schools were performing on the English SOLs. Level one in green marks that they are meeting the state-set benchmarks; level two in yellow means they are approaching the benchmarks; and

level three in red means that they have significant improvement necessary to meet the state benchmark.

All teachers in the three schools who taught classes that had an English SOL assessment requirement were contacted via email and asked if they were willing to participate in the study. Six teachers self-selected into the study. One elementary teacher, one general education middle school teacher, three middle school special educators, and one high school teacher agreed to participate. Each participant was given a pseudonym. In addition, two individuals at the Department of Education were interviewed simultaneously to gain perspective about the policies-as-intended at the state level. Demographic information is organized in Table 12 for easy reference.

Each of the participants was interviewed for about an hour, totaling seven hours of interview data. All interviews were conducted via Zoom videoconferencing software due to the COVID-19 pandemic. In addition, observational data was collected for one class period during the project-based learning (PBL) class in which high school teacher taught, totaling approximately 90 minutes. Three additional meetings were observed: a division Facebook Live meeting held by the superintendent and administrators of the division and two virtual State Board of Education meetings, totaling approximately 5 hours of observational data. Although this study was initially designed to have additional observational data, data collection began in March 2020 and was significantly impacted by the COVID-19 pandemic. This was a significant limitation of the study. Documents from the district and state-level that were relevant to the meetings and policy documents about the SOLs and Profile of a Virginia Graduate were also collected.

Figure 1

Characteristics of Sample Schools
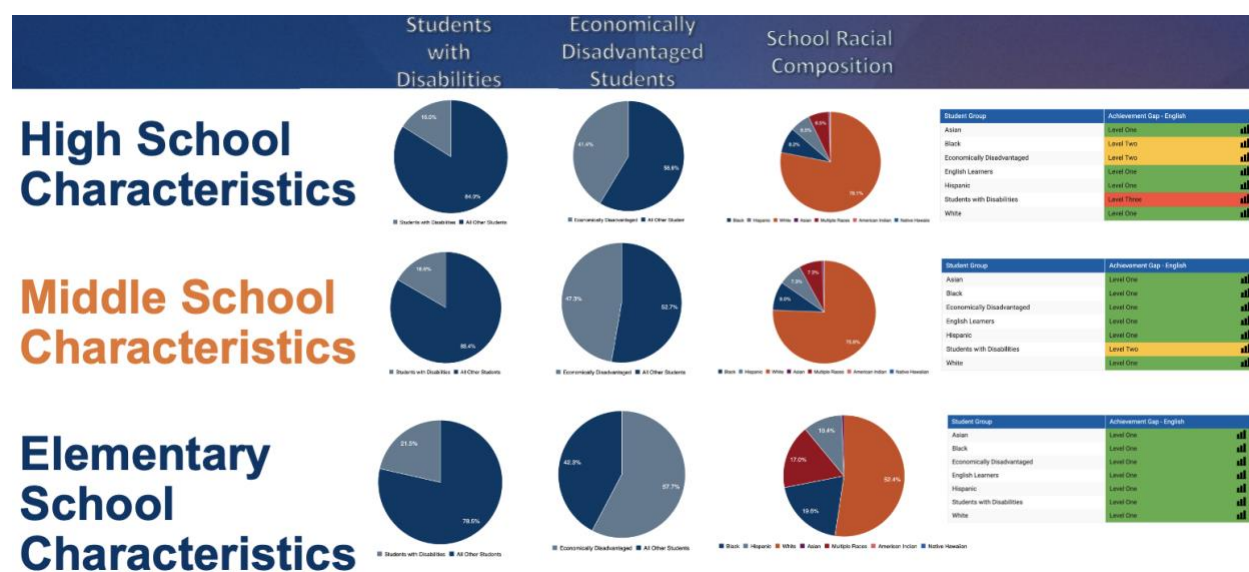
Note. Level One – School is performing at or above the state standard for the indicator; Level Two – School is performing near the state standard for the indicator or improving; Level 3 – School is performing below the state standard for the indicator. Demographic data and graphics from schoolquality.virginia.gov.

Table 12

Participants and Demographic Information

Pseudonym	Placement	Years of Experience
Ms. Sphinx	Elementary – 4th grade	22
Mr. Lars	Project-based learning middle school within a school – 8th grade, former career/technical teacher at the high school level	15
Ms. Cardinal	Middle school – 7th grade special education inclusion classes	8
Ms. Mayo	Middle school – 8th grade special education inclusion classes	24
Ms. Evergreen	Middle School – 8th grade special education inclusion classes	6
Ms. Moses	Project-based learning and traditional classroom – 10th grade	13
Ms. Orsen	Liaison for the Board of Education	3.5
Ms. Rue	Office of Policy	1.5

Analysis of Data

      As suggested by Erickson (1986), analysis was conducted through analytic induction. The data corpus, including interviews, observations, and documents, were used to develop an understanding about teacher interpretation and implementation of state policies. First, I repeatedly read through all the data collected and wrote analytic memos about each data source. From this, patterns and codes began to be evident as data sources were looked at holistically, compared, and contrasted. From these memos, I was able to create draft analytic assertions. Assertions are the conclusions drawn from the data through this process of analytic induction. Once draft assertions were formed, the data corpus was searched for evidentiary warrant, elements that confirmed and/or disconfirmed the assertions. Since analytic induction is an iterative process, as assertions began to come to light, teachers who mentioned topics that fit within the assertions were probed during interviews to help identify evidence that confirmed or disconfirmed the assertions. When disconfirming evidence was found, it was used to adjust the assertions so that the assertion fit the data. In the results section the assertions are presented along with evidence from the data to support assertions is offered in the form of quotes and vignettes.

Limitations

      One limitation of this study put in place by COVID-19 was a limit on the kinds of data collected. Although some data was collected through interviews, observations, and material culture, none of the data sources were as plentiful as originally intended, especially the observational data. To help offset this limitation, interviews were carefully studied for confirming and disconfirming evidence. Disconfirming evidence was thoroughly analyzed and used to tighten and change assertions. In addition, thorough evidence is presented for each

assertion so that the reader may make the determination for him or herself if the interpretation has been faulty.

This study was designed as a case study. Flyvbjerg (2001) suggests that the most meaningful social science work is done at the local level. Although the context of the study is specific to one school district, concepts illuminated have the potential to be beneficial to other districts.

Researcher as Instrument

I have elementary and middle school experience in teaching. I believe this perspective was beneficial as a researcher because it allowed me to better understand what I was seeing in class observations and helped me intuit helpful follow-up questions to ask during interviews. Practicing reflexivity allowed me to recognize my biases and conduct this research with a helpful blend of emic and etic perspectives.

Throughout the process, I ensured participants that they were the experts and that I was there to learn from them in order to help negate any intimidation or fear that may be experienced by participants. There were several times during probes where I also reminded participants that I was not evaluating them in any way, but rather wanting to better understand their perspectives in order to encourage them to share freely about their opinions and practices.

<div align="center">

Results

</div>

The assertions found in this study are as follows:

1. Time restrictions meant that teachers found the Standards of Learning and 5 C's difficult to implement concurrently.

2. Teachers appreciated the Standards of Learning as a guide but did not like the test, especially in its current form, because of the way it distilled a year's learning into a single performance.

3. Rigid instructional mandates from school administrators and 5 C's implementation were inversely related

4. Prioritization of the SOLs versus the 5 C's was influenced by the varying power of stakeholders, with administrators often shaping actions at the local level

5. DOE employees have convergent and divergent perspectives when compared to teachers

Assertion 1 – Time restrictions meant that teachers found the Standards of Learning and 5 C's difficult to implement concurrently.

All teachers interviewed responded negatively when asked to consider the concurrent implementation of the SOLs and the 5 C's. From elementary to secondary, all teachers agreed that the two policies, as implemented, were not supportive of each other. When asked how teachers saw the two policies interacting or reinforcing in their classroom, Ms. Mayo simply said, "I don't" while Ms. Moses said, "They're not."

Ms. Sphinx pointed out that the issue was not with the standards themselves, but with the test. She explained, "I think they would interact better if there wasn't a test as the end result." Mr. Lars felt it was difficult to merge the two frameworks in practice and doubted his own ability to accomplish this. He said that core teachers probably do a great job with this, but "I have troubles, in my own head, designing a lesson where they would be together." In actuality, the core teachers also had difficulty with designing a lesson that integrated both policy frameworks.

Ms. Cardinal explained a tension and conflict between the two policies:

They don't seem to complement each other very well. In my opinion, the 5 C's…. If I'm

using them the way that I would like to, you can kind of be creative in them. You can –

It's higher-level thinking. SOLs, you know, you just gotta hit these points, hit these points,

hit these points.

In the key phrase "if I'm using them the way I would like to," Ms. Cardinal indicated that a form

of the 5 C's could be used, but that the conflict with the Standards of Learning prevented those 5

C's skills from being implemented in the way she believed they were intended to be. Other

teachers agreed with this point. Several teachers noted they could use partner work or label

something as addressing a 5 C, when in reality the skill was not a focus and was not being taught

in a way that gave the skill the attention it deserved, meaning the rote memorization was

prioritized over the higher order collaboration skill.

The Profile of a Virginia Graduate's 5 C's are an additive policy that assumes teachers

can do it all without removing any expectations that were already in place. Ms. Orsen, with the

Board of Education, acknowledges this limitation, saying, "you only have so much time in the

day to get through everything." With the Standards of Learning already being familiar to

teachers and assessed with high-stakes testing, it is not surprising that if teachers find

simultaneous implementation difficult, the SOLs get prioritized.

Assertion 2 – Teachers appreciated the Standards of Learning as a guide but did not like

the test, especially in its current form, because of the way it distilled a year's learning into a

single performance.

When asked what they would change about the Standards of Learning, every teacher

interviewed made reference to wanting to change the test. Some would get rid of it altogether,

others wanted to change the way it was used, but no one thought it should be used in the same

way that it is currently being used, with a year's worth of content tested with one high-stakes test used for accountability at the student, teacher, school, and district level. As Ms. Moses put it, the existing system encourages what she referred to as the "sit, get, spit, forget model" in which students learn the material only to regurgitate it on a test. Ms. Mayo succinctly stated what many teachers indicated about the Standards of Learning:

> *I do think it's a good idea to have Standards of Learning, you need standards. However, I think teachers should be given the leeway to decide how they want to teach it, or maybe they want to do a project... instead of giving like a common assessment at the end.*

This leeway that she mentions is totally usurped by the test. Even when things show up in the Standards of Learning, they only become a focus when they also are known to be a tested standard as explained by Ms. Cardinal:

> *Some of the standards are good, and they don't even get focused on, like, you know, there are standards that don't get tested [She air quotes these three words to indicate that someone has said this to her]. So, we don't do them.*

Here, an issue with accountability is highlighted again. Even within the Standards of Learning, when certain parts are associated with high-stakes consequences and others aren't, the focus often shifts to the component with the consequences, despite the preference of the teachers.

Ms. Sphinx also indicated her frustration with the inauthenticity of the test. She stated, "Well, I guess I don't see a great big need for the end of the year SOL test." When asked what she would prefer, she said:

> *Some sort of real-world performance task where they actually have to apply what they know, and talk about what they know, and-- You know, when you get to go to a job interview, you're not going to sit down and take a test. You're going to have to have a*

conversation with another person and they're going to have to understand that you know

the concepts of what they're asking you to do.

Under the model she envisions, the incompatibility of the SOLs as standards and the 5 C's would be lessened. If allowed to have performance assessments, students would be able to address the skill sets within the standards in ways that encouraged each of the 5 C's.

Teachers' responses to the SOLs can help us to understand one of the reasons the Profile of a Virginia Graduate was introduced, to make the connection between school and the "real" world. However, as stated before, the addition of a policy does not necessarily mean that the desired connection between school and post-secondary life is being forged.

Assertion 3 – Rigid instructional mandates from school administrators and 5 C's implementation were inversely related.

The context of the classroom had a significant impact on the implementation of the policies. Two of the teachers who chose to participate in this study were part of two different school-within-a school models that focused on project-based learning (PBL). These teachers provided a glimpse into a few special classrooms across the division that serve as anything but the norm. In these classes, there were fewer rigid expectations about how to teach placed on the teachers by administrators. In the PBL schools, the classes were expected to be taught through large unit projects that were used as assessments versus the more traditional lecture or activity-based learning, assessed by quizzes and tests. In these PBL classrooms, teachers had freedom in their lesson design and implementation of the projects.

In explaining the difference between the two settings, Mrs. Moses stated:

This dirty little secret here, because [in the PBL classroom] I don't have to submit

lesson plans in a certain format, with certain vocabulary, on a certain day, for certain

people. I've got some flexibility. Now for my [traditional class], it's different… more lockstep, I have to be doing what [she says in a robotic voice hitting her hand down with every word] every other English 10 class is doing.

Vignette

To understand how the difference plays out, consider Mrs. Moses' high school English and science co-taught project-based class, in which they presented a new project. Students listened eagerly to what their next challenge would be. Students were tasked with a research project about a local natural preserve. Their job was to research the preserve and create a presentation that they then presented at the local agricultural fair.

Students worked in groups, practicing collaboration and communication as they determined what information their audience needed and when sections of the project were "good enough" before moving on to the next. About one-third of the way through the class, as students were working, one of the teachers clapped her hands, "Stop! Right now, score yourself on your rubric. Where are you?" Each student had a sheet of paper with three copies of the exact same rubric on it. The rubric had four categories: emerging, developing, proficient, and exemplary. These categories described their current collaboration and contribution within their group. On the top rubric, the students circled one of the categories and wrote a very brief explanation why that was their choice. With no commentary, as soon as students filled in their rubric, they got back to work.

An approximately equal time later, the process repeated. "Okay, stop! Where are you now?" By the end of the class, every student had moved up at least a box on their rubric.

Through this process, students were able to identify early in the class period that they were not where they needed to be. They could see that they would be self-assessing again in the

future, so they rose to challenge, and by the end of the class period, everyone was at proficient or exemplary.

Every few days, the teachers held a debrief about this process. They would ask, "Look at your pattern here. All right, what is behind that pattern?" From this, the students expressed that it often took them some time to settle into their assignments. This self-reflection helped the students to better understand their own learning needs, and it helped the teachers to better facilitate the class in order to meet those needs. As the whole class convened to self-reflect on their collaboration skills, they identified obstacles and problem solved ways to overcome them.

As the month progressed, the students took hikes up the local nature preserve experiencing first-hand the wonders it had hidden within it. When they stumbled upon a question that they didn't know the answer to for their project, they participated in authentic research, scouring the web to find the answers to their questions. Ms. Moses distilled the secret to the success of this type of project, "When they're invested in it, then they pay more attention."

In addition, students had the opportunity for their strengths to shine. As the groups worked together, their individual strengths and skill sets complemented each other as they relied on the group to fill in for each other's deficits. This allowed for authentic collaboration that more closely mimics team projects outside of school, where individuals are valuable to a team because of their strengths and aren't necessarily punished for their weaknesses. For many of the groups, the hard work paid off in more ways than one. Several groups were asked to present the findings of their research at a local scientific conference after the agricultural fair.

Later in the year, once stay-at-home orders were in place, the students from this class were able to use the 5 C's as they worked on their capstone project for the class. When they were

first presented with the assignment they expressed, "We don't know how to use all these

programs! How are we supposed to do this project?"

Ms. Moses rebutted, "Bingo! That's the point…One of the skills that people need to have

is they need to learn how to learn. If you think this will never happen in real life, welcome." She

mimics her hands in an arch shape as if saying "ta da!" "Welcome to the moment because every

single teacher you have has spent the last two weeks doing exactly what you're doing right

now." Even if they had doubt, these students were able to take the skills they'd developed during

the year, and they met this challenge. During online class meetings, the students identified their

problems, asked for what they needed, and advocated for themselves.

In another example of this phenomena, Ms. Cardinal compared her teaching experience in a private school setting to her current public school setting.

Something I really enjoyed being able to do… I could go down rabbit trails when they

were interested in something, or I could let their curiosity dictate the lesson that day. You

can't do that with SOLs because you have to – You have to stick to it… because that test

is coming. You have that sense of urgency and stress and then the kids feel that too,

obviously.

She had shared how loose the requirements were in the private school. Even though she stated that she is a public school advocate, she missed teaching without the unyielding requirements she faced in her public middle school setting.

Both Ms. Mayo and Ms. Sphinx stated that they thought the 5 C's were good "in theory." When asked what they would change about them, both simply stated that they wished they had the freedom to implement them. While Standards of Learning were expected to be documented on lesson plans and assessed, it was up to the individual teacher to note progression on the 5 C's

without external accountability. The rigidity of expectations within one policy and the flexibility in the other led to the SOLs dominating the time and energy of the teachers, with the 5 C's serving more as an afterthought.

Assertion 4 – Prioritization of the SOLs versus the 5 C's was influenced by the varying power of stakeholders, with administrators often shaping actions at the local level.

All but one teacher said that they believed the skills taught within the 5 C's were more important than the skills focused on within the Standards of Learning. They referred to the 5 C's as life skills and mentioned that they were needed no matter which career or job a student chose. Flyvbjerg (2001) acknowledges that power is at the core of all human interactions. When carefully investigating what at first seemed like disconfirming evidence, the one teacher who did not definitively say that he thought the 5 C's were more important to students in the long-term, made a nod to power in his statement. Mr. Lars stated:

> *I think that may have do with why I prefer SOLs. It's more of a comfort level as opposed*
> *to I think one is better than the other. I think if I got switched into 5 C's, and we were*
> *pushing that a little more heavily. I would probably be answering this differently.*

From this quote, we can see that it is not the superiority of the skills assessed by the SOLs that Mr. Lars prefers, but instead he implicitly acknowledges that the policy that is prioritized by the school and system, the one that has been in place and that he knows, because it is discussed or "pushed" more heavily, is the one he prefers.

Teachers asserted that school administrators discussed the SOLs to a much greater degree than the 5 C's. Ms. Mayo explains:

That's all we hear about: SOLs, and if we don't pass and we're not accredited, the state

is going to come and hang out with us all [she exaggerates this word drawing out the all]

day long and tell us what to do.

The quotes from Mr. Lars and Ms. Mayo illustrated how power is at play in the classroom. Teachers did not feel they have the freedom to implement the 5 C's because of the focus on the Standards of Learning and on the test. Teachers feel they must focus in their classrooms on what administrators thinks is important, and administrators choose what to focus on based on what they believe is required by the state.

Many teachers spoke about the importance of a growth mindset, the focus on growth as opposed to achieving a particular standard at a particular time. These teachers all expressed how they wished that growth could be accounted for by the Standards of Learning. Those who didn't explicitly speak to growth, implicitly acknowledged a wish for a measurement of growth as opposed to a summative assessment with the Standards of Learning.

Interestingly, according to the Standards of Accreditation, a policy shift toward growth was implemented in 2017 for the 2018-2019 school year (Virginia Department of Education, 2019). For reading and math, students who showed significant growth on the SOL tests were counted toward the pass rate for the school. When teachers were probed about this, some had heard of the concept, but only Ms. Mayo seemed aware that this policy was already in effect. She indicated that although this was the case for reading, since writing wasn't assessed every year, this growth model did not help in that subject area.

In addition, in November 2019, the Virginia BOE changed the evaluation model for teachers. Where formerly assessment scores accounted for 40% of a teacher's evaluation, in the new model SOL scores were equal with the other components of the evaluation (Observation,

BOE meeting, November 13, 2019). Teachers still indicated, however, that they felt that their worth at the school was very much tied to their SOL scores, as indicated by Ms. Evergreen:

> *I don't think anybody's going to take their eyes off the SOLs until they're not key to keeping your job, to keeping your school…it would be impossible for teachers to take their eye off that ball until they knew that all of those things weren't tied to SOL scores anymore.*

Regardless of how much of their evaluation was tied to their scores, teachers felt that they were responsible to what their administrators requested, which included high test scores. An illustrative example of this is seen in Ms. Cardinal's words as she speaks about the pressures she feels under her administration and how they have the power to change the way she feels:

> *I think that if administration… could … say, "Hey, you know, this is no longer. It's okay to fail. It's okay to have a different metric."*

Until those with supervisory power give permission for teachers to try new ways of teaching that might be prone to a few rounds of trial and error, teachers are too scared to try for the worry of being burned by low test scores and scorned by their administration. Therefore, even for teachers who see the 5 C's as a more important skill set than the SOLs, changing instruction to be more aligned with the 5 C's than the SOLs won't happen without administrative push and permission.

Assertion 5 – DoE employees have convergent and divergent perspectives when compared to teachers.

Both employees of the DOE interviewed understood that SOLs were largely thought of as synonymous with the test, but they indicated their belief that there was flexibility associated with the standards. Ms. Orsen, who is DOE staff person who works with the BOE, spoke about how the 5 C's were being assessed:

I think it's going to vary across the state how that is done because we don't have specific

assessment attached to these five skills. This is really where that, you know, performance

based assessments and project-based learning can come into play in a student's

education career.

While she realizes that in PBL settings, the skills are being assessed and utilized, the down-side

of the lack of "specific assessment" isn't addressed. In the case study of this district, in the

traditional classroom, this lack of *specific* assessment often resulted in *no* assessment.

Another divergent DOE perspective was that the SOLs and 5 C's were "two pieces of the same

puzzle" (Ms. Rue). This of course is a vastly different response from that mentioned by teachers

in assertion one. Ms. Orsen adds:

The standardized assessment movement has changed over the last decade, whereas in the

90s, [standardized assessment] was all. [Now] we know that [standardized assessment]

may not be the best way to actually assess the students' knowledge or their skills. So,

while we may not have specific assessments for the five C's, there's also a way to assess

those five C's as you're assessing that knowledge in a variety of ways. So, it can really

allow us to kind of rethink how we do assessments and we're starting to do more of that

as we implement The Profile of a Virginia Graduate.

So, while teachers interpreted the 5 C's as an additive policy, evidence here supports that the 5

C's are intended to give more flexibility to classroom teachers. Ms. Orsen and Ms. Rue indicated

the changes in evaluation for teachers that de-emphasized standardized test scores to support

their claims, but as stated in assertion 4, teachers don't seem to be yet feeling or enacting the

change in this intention.

When asked about implementation of the SOLs, unlike teachers, Ms. Orsen and Ms. Rue agreed that they wouldn't change the actual implementation, but rather how administrators and teachers *interpret* the SOLs, as Ms. Orsen stated:

> *I don't know that I would change anything about how they're implemented, what I would*
>
> *change is the misperception that the SOLs and passing the SOLs is the be all-end all.*

In contrast, all the teachers interviewed indicated that there were big changes they would make, another clear divergence from intention and practice.

It is interesting that the teacher comments about the importance of the SOLs are contradicted by what the DOE staff states. While teachers see accountability as of utmost importance, the staff indicates here that the scores are only seen as part of a greater picture. Ms. Rue does acknowledge, however, the role of administration in this shift stating, "I wish that…school leaders wouldn't put so much pressure on, especially some of our younger grades, to pass SOLs." Here she indicates her understanding of the importance of school leadership in policy transformation, as indicated by Hall and McGinty (1997).

From these perspectives, we can tell that the wishes or intentions of the DOE and teachers aren't too misaligned; however, their conceptions of practice are *not* aligned. They both agree, however, that administration has a key role to play in the state's culture change.

<p style="text-align:center">Discussion</p>

Flyvbjerg (2001) suggests that social science research should be working with the question "What should we do?" in mind (pp. 136-7). From the perspectives of teachers as professionals, it seems clear that what we should do and what we are doing are not aligned. Rather than relying on their own judgement and values, teachers made sense of the state policies

after they were interpreted by the administration. If administrators gave teachers the freedom to implement the 5 C's, such as in the PBL classrooms, teachers did.

Secondly, when considering the differences between policy intentions and outcomes, accountability was a key player. Decisions about accountability directly impacted policies' outcomes. Due to time constraints, teachers were not able to teach both skill sets, those dictated by the SOLs and 5 C's, well. While many teachers saw the 5 C's as more important, the skills were given much less emphasis than the rote tested skills of the SOLs. While the Profile of a Virginia Graduate may have been developed with great intentions, the policy just isn't enacted in classrooms in the same way as those that are signaled to "matter" to students, teachers, and schools through the consequences in place around them.

This has important significance for policy makers. As the DOE has begun to lessen the number of standardized tests, they have essentially given more power to the areas that are still tested. Without a total shift in culture, it will be difficult for those policies without evaluation in place, like the 5 C's, to hold their weight against those that have reigned for so long, such as the SOLs.

Conclusion

The lessons of the pandemic are clear: School systems and teachers have adapted and continued to teach. Students have continued to learn, even without a standardized test in place. With new experiences behind us and new possibilities before us, teachers, administrators, and policymakers have an opportunity to implement changes.

As we answer Flyvbjerg's "What should we do?" (pp. 136-7), there is an opportunity to learn from the assertions developed from the words of the participants of this project. When given the right administrative support, teachers can cover both content and soft skills, and

students can learn them; however, while longstanding systems of assessment and accountability exist, they prevail. As we move forward, out of the pandemic and back into classrooms across America, may the people in power trust the professional teachers to do such.

References

Braun, H. (2004). Reconsidering the impact of high-stakes testing. *Education Policy Analysis Archives*, *12*(0), 1.

Erickson, F. (1986). *Qualitative methods in research on teaching* (pp. 119-62). Institute for Research on Teaching.

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press.

Grissmer, D., & Flanagan, A. (1998). Exploring rapid achievement gains in North Carolina and Texas. *Lessons from the States*. Washington, DC: National Education Goals Panel.

Guba, E. G. (1984). The Effect of Definitions of" Policy" on the Nature and Outcomes of Policy Analysis. *Educational Leadership*, 42(2), 63-70.

Hall, P. M., & McGinty, P. J. (1997). Policy as the Transformation of Intentions: Producing Program from Statute. *The Sociological Quarterly*, *38*(3), 439-467.

Nation's Report Card. (2019). Results from the 2019 mathematics and reading assessments. Retrieved from

https://www.nationsreportcard.gov/mathematics/supportive_files/2019_infographic.pdf

Newmann, F. M., King, M. B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review*, 67, 41–74.

Nichols, S. L., Glass, G. V., & Berliner, D.C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives*, *14*(0), 1.

Nichols, S. L., Glass, G. V., & Berliner, D.C. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education Policy Analysis Archives*, *20*(20), 1.

Virginia Department of Education. (2019). Guidance document: Governing certain provisions of the regulations establishing standards for accrediting public schools in Virginia. Retrieved from http://www.doe.virginia.gov/boe/accreditation/index.shtml

Virginia Department of Education. (2020). *Profile of a Virginia Graduate*. Retrieved from http://doe.virginia.gov/instruction/graduation/profile-grad/index.shtml

Paper Three: What Does Success Look Like? Teachers' Realities and Reimaginings in K-12 Schools

Sarah Beach & Helen H. Min

The unprecedented outcomes of the global coronavirus pandemic challenged the barriers between professional and personal lives, and this sudden departure from the status quo forced citizens to consider the systems surrounding them in new ways. One such system was the U.S. public education system, including its organization around standards, assessment, and accountability. This study examines how the pandemic challenged existing boundaries in the education system and how lessons learned from 2020-2021 may help us reimagine the success and wellbeing of teachers and students. These considerations will be filtered through the lens of the standards, assessment, and accountability system to explore the ways the system impacts teachers and students and influences teachers' visions of a reformed educational imagination.

For too long, systems of education have relied on metrics such as standardized tests as ways to evaluate both students and teachers within the education system. Reynolds and Howell (2010) indicated that one of the "traps" in conventional thinking, as opposed to systems thinking, is the "focus on outcomes (and thus only on what can be measured)" (pp. 5-6). Systems thinking encourages holistic rather than outcomes-oriented thinking, and thus the United States (U.S.) education system would benefit from reorienting the current understanding of success. Utilizing data collected from teachers during the pandemic, the findings of this research indicate that teachers and students yearn to be seen as whole people, and that redefining success holds the potential to transform the education system for increased sustainability in a post-pandemic world. The study illuminated how traditional assessment and accountability policies acted as

barriers to the redefinition of success and engaging a new and more expansive view of teacher and student success.

Literature Review

The context of the pandemic has shifted roles and responsibilities within the education system, adding additional stress and constraints. Teachers are switching careers, schools are struggling to support teachers, and students are left in the flux of navigating uncertainties.

Student Success

Since the beginning of the No Child Left Behind policy, educators and academics have investigated the competencies students need for the 21st century (e.g., Partnership for 21st Century Skills) along with the accountability metrics measuring student success (Kay & Boss, 2021). These have focused primarily on standardized test scores, with some movement toward including additional "soft skills" such as communication, creativity, collaboration, and critical thinking developed by Partnership for 21st Century Skills, "Seven Survival Skills" (Wagner, 2010), global competencies (Reimers & Chung, 2019; Zhao, 2012), and digital citizenship competencies (Tamayo, 2016).

Unfortunately, within the education system, students are often quantified and dehumanized through test scores. Students' varied funds of knowledge (Moll et al., 1992) are not tested or valued when success is limited to the results of one end-of-year standardized assessment. Teachers are then held responsible for this success, another act of dehumanization of students. If we think about this systemically, it is as if students are products within a factory, to be produced and measured for quality assurance. When students don't meet the mark, the workers, teachers, are held responsible.

Teacher Mass Exodus

In 2021, nearly one in four teachers wanted to leave the classroom, compared to the pre-pandemic average of 16% (Zamarro et al., 2021). High rates of stress associated with the teaching occupation (Gallup, 2014) have been exacerbated by staffing shortages which have contributed to increased teacher workloads with decreased support (Lieberman, 2021), multiple changes to instructional demands (Zamarro et al., 2021), and challenges to the personal lives of teachers and students due to the pandemic (Steiner & Woo, 2021). The confines of a system not already imbued with flexibility have led to increased overburdening of teachers and students. Without meeting the complex needs of both teachers and students, teachers in the educational system suffer from increased burnout and turnover. In the long term, this may contribute to a lack of consistent and qualified teachers, poor teacher performance (Greenberg et al., 2016), and negative impacts on students' academic outcomes and social adjustment (Oberle & Schonert-Reichl, 2016). Addressing this stress is essential for both teacher and student success.

Conceptual Framework

Historical Perspective and Systems Thinking

While the context of the pandemic is imperative in understanding the current challenges facing the U.S. education system, considering the historical perspective of how the system came to be is also beneficial. While modern technology, thought, and expression have continued to grow and diversify at exponential rates, the education system is still largely rooted in ideas from the advent of industrial times. As machines were being developed, reductionistic thought thrived (Ackoff, 1974). Society sought ways to simplify and make life easier. After the Industrial Revolution, society began expanding again, thinking of ways to open minds and consider multiple ways of knowing; however, the education system has continued to prioritize the values

of efficiency and competition, making little room to incorporate expansionistic thinking (Jennings, 2020). The development of the technology of assessment has paralleled these historical trends and have contributed to this problem.

Systems thinking considers not just individual parts that make up a whole system, but rather the system itself as a unit and the roles that each of its sub-parts play within that whole (Ackoff, 1974). Considering the components of a system addresses problem solving from a macro perspective, seeing not just single, isolated problems, but what Ackoff (1974) coins as a *mess* of problems that are interrelated and complex, creating other problems as they interact with the system (p. 21).

Historically, much of the burden of the flaws of the system have been placed upon teachers (e.g., Ackoff, 1974; Guggenheim et al., 2011; National Commission on Excellence in Education, 1983). Despite acknowledgement of the intertangled mess of problems, teachers are often burdened with additional responsibilities or villainized for the shortcomings of schools without recognition that they are only one subpart of the system. This study utilizes a frame that recognizes the complexity of the web of problems that leave teachers feeling disempowered.

Critical Systems Heuristics

In this study, we employed the lens of Critical Systems Heuristics (CSH; Ulrich, 1996; Ulrich & Reynolds, 2010) to understand teachers and their work within the education system at large, with particular attention to systems of assessment and accountability. Attending to power dynamics within systems of education, this framework of systems thinking is useful in analyzing data through the lens of the teachers involved in the system.

Ulrich and Reynolds (2010) suggest that there are three major reasons for using CSH: to better understand the big picture, to promote better understanding from different perspectives,

and to aid in reflection to promote change of the system being studied, in this case the assessment and accountability system in place in public schools. Through CSH, people can understand that the way they perceive and navigate the world likely differs from others involved within a system. The CSH framework allows these differences to be uncovered and interrogated so that sources of contention and misunderstandings rooted in perception can be alleviated.

This study's application of the CSH conceptual framework allows for the comparison of the system *as it is* versus how it *ought to be* from the perspectives of teachers. A *normative* educational system observes the perceived realities of the system in the present status quo, while the *reimagined* educational system envisions what the system should be. The juxtaposition between the normative and reimagined systems allows a consideration of the complicated *mess* of the current education system to be explored as means "of coping with the circumstances as best one can" (Reynolds & Howell, 2010, p.5). This *ought to be* system is tied to critical reimagination literature (Kelley, 2022; Spaulding et al., 2009). Kelley (2022) asserts that "without new visions we don't know what to build, only what to knock down" (p.xii). This provides the purpose for examining the ways in which accountability and assessment policies have served as barriers to systems changes and the new futures we want to move toward.

Critical Care Pedagogy

In considerations of reimagined systems, the conceptual framework of critical care pedagogy is employed. Critical care pedagogy is an approach to education that emphasizes the importance of caring relationships between teachers and students, recognizing that caring is a complex and culturally specific concept that must be understood in the context of different social, economic, and cultural factors (Antrop-Gonzalez & De Jesus, 2006, Johnson, 2014; Killam & Camargo-Plaza, 2022; Love, 2019; Rolón-Dow, 2005). This pedagogical framework

draws from three theories of caring: teacher caring theory, caring community theory, and difference theory. Teacher caring theory posits a causal relationship between teachers' caring behavior and student success (McKamey, 2004). The theory of a caring community recognizes that schools and communities have both the ability and responsibility to create supportive environments for students. Difference theory acknowledges the myriad definitions of caring stem from social, ethnic, class, and gender groups and asserts that schools that incorporate these differences better foster senses of belonging (McKamey, 2004).

Importantly, critical care pedagogy aligns with difference theory, critiquing the color-evasive, sometimes referred to as colorblind, assumption in White feminist notions of caring, which often is characterized by lowering academic expectations out of an emotional response of pity for students' social circumstances (Katz, 1999; Thompson, 1998). Instead, difference scholars suggest that caring has existed within the sociocultural, gendered, and economic contexts for disenfranchised communities as a public undertaking as well as a private or semi-private concern (Thompson, 1998). Rather than forcing students to conform to the teachers' expectations of caring, Valenzuela (1999) argues that conceptualizations of caring in education "must more explicitly challenge the notion that assimilation is a neutral process so that cultural and language affirming curricula may be set into motion" (p. 25). This approach emphasizes building on students' cultural and linguistic knowledge and heritage to bolster the cultural and linguistic assets of youth, rather than subtracting these identities from them to their social and academic detriment.

The framework of critical care pedagogy is bolstered by ongoing research. In a 2021 survey of 1,379 K-12 parents, teachers, and administrators across the United States, Instructure (2022) found that 97% of administrators, 98% of teachers, and 90% of parents identified student-

teacher relationships as one of the most impactful social-emotional factors for student success. This highlights the value of teachers in enhancing the learning experiences of students and emphasizes the need for change in the way student success is assessed. Thus, this requires professional development opportunities that enable teachers to adopt innovative methods to evaluate and give constructive feedback to their students. Additionally, teachers need to acquire skills through professional development that facilitate their understanding of the diverse social, economic, and cultural backgrounds that students bring to their classrooms. It is also crucial that teachers learn to develop positive relationships with their students that affirm their unique identities both relationally and in the curriculum.

Methods

Using CSH as a conceptual framework, empirical data was collected and analyzed to help better understand how teachers are thinking about success within the system for themselves and their students. The following research questions guided the investigation:

1. How do teachers make meaning of success for themselves?

2. How do teachers make meaning of success for their students?

3. How do teachers want to define success?

4. How do state systems of assessment and accountability influence teacher meaning making of the system?

The data for this study were collected in two different settings. One set of data included 24 lead teachers in K-12 classrooms across the United States collected in April 2021. Teachers who completed a demographic survey and interview consent form through a flier link posted on social media (i.e., Facebook, Instagram, and Twitter) were contacted by email. To maximize variation in the sample for this chapter, we purposefully sampled nine interviews from this larger

dataset. Of the nine, five teachers identify as male, four as female, five as White, two as

Hispanic/Latinx, two as Asian, three at the elementary level, three at the middle school level, and

three at the high school level.

The other data were collected in March through May 2020 and included six teachers

selected through convenience sampling within a single Virginia school division: one at the

elementary level, four at the middle school level, and one at the high school level. Of these

teachers, five were White and female, and one was Black and male.

While neither of these samples are representative of the U.S. teaching population as a

whole, together they allow for a deeper understanding of teachers' meaning making by seeing

where the definitions of success converge and diverge among teachers in different settings with

different experience levels.

To analyze the data, deductive coding was applied using the conceptual framework. First,

the data corpus was read and coded into teacher data and student data. While all interviews were

given by teachers, when data specifically referred to students or their success, this was coded as

student data. This allowed looking at meaning making about teacher success and student success

separately. Next, both student and teacher data were coded using the codes "success is" and

"success ought to be," codes determined a priori from the CSH framework. The researchers

coded the first interview together to set up procedure and discuss alignment, then the remaining

interviews were coded individually. After coding, the researchers read through all of the coded

data for confirmation and discussed any codes which had been marked as unclear. While coding,

the themes of "relationships," "burn out," and "barriers" became clear, and codes were added to

distinguish that data.

Next, the data were analyzed using thematic analysis. The coded data were read and connections were made to determine themes within how teachers defined success as is and in reimagined ways. The data were read again for confirming and disconfirming evidence. Finally, the authors read through the preliminary themes to merge and revise themes. The final themes are presented in the following sections.

Results

What Teacher Success Is

Within the system as is, teachers defined their own success in three major ways: helping students achieve high test scores, completing all parts of their job, and managing their classrooms. These three components were essential not because teachers necessarily saw them as most important, but because they believed these were the indicators of success that were incentivized by the system, with the testing and accountability system serving as a lens through which the other components were viewed.

Test Scores

Teachers reported that their success is measured by the performance of their students on standardized assessments, district benchmarks, and teacher created exams. Standardized assessments include exams created by state departments of education or national testing agencies that measure whether students are performing at their grade level based on standards detailed by the state or national agency. District benchmarks often use released items from these standardized tests and are aligned with the same standards. Teacher created tests may measure students' performance on a subject at the beginning of the school year through diagnostic tests, assess students' progress throughout each lesson through summative tests, mark students'

mastery of a learning objective or standard through formative assessments, and measure

students' overall understanding of a subject through summative exams.

This emphasis on the outcome of student test scores, often synonymous with student

achievement, became a central priority in measuring teachers' success because of what the

accountability systems, and thus administration and school districts, communicated directly,

nonverbally, and through teachers' pay systems. One teacher reflecting on their first few years in

the profession remarked that they were seen as successful if they followed the curriculum given

to teachers by the school districts and followed the progression of the assigned textbooks.

Teachers reported that there is a clear gap between what teachers think will be successful for

students and how the system measures success. One teacher explained:

> *Well, it just feels cruel. What we do to core teachers. In terms of setting up evidence that*
> *they've taught, and in the end, no matter what other stuff is done inside the classrooms,*
> *we kind of come down to the end and go, "Okay, what are those [state standardized test]*
> *scores?"[1]*

Here, the teacher deplores the standardized testing system on which "core teachers," teachers

who teach tested subjects like English and math, are gauged for their success. The teacher

explained that other actions that the teacher conducts for their students and their learning are

irrelevant if the students are not tested in that content area.

On one hand, while teachers knew that their value as teachers relied more heavily on their

relationships with students and other intangible aspects (e.g., classroom culture, students' love of

learning, students' practice of social and emotional skills), they were told that their jobs were

---

[1] Quotes have been edited slightly for readability, editing out fillers such as repeated words, "um," and "you know."

dependent on the state standardized test scores of students in their class. This inevitably became a big stressor for teachers:

> *That's what we've been told over and over, like basically our jobs are tied to our scores. Our accolades are tied to our scores. We have been told whether we are a good teacher or not is tied to our scores.*

Similar to other teachers, this teacher highlighted the repetitive emphasis placed on students' academic achievement as measured by the standardized assessments. Through this messaging, state, district, and building leaders communicated a connection between teachers' occupational worth, professional efficacy, and financial stability with students' scores on the end-of-year exams. Details relevant to the assessment, such as the teachers' ability to prepare students for the test or students' growth on the exam based on their baseline scores, were often not accounted for and did not factor into the measure of teachers' success as it related to their job security, professional achievement, or financial reward.

The pressure to prepare students to score highly on the standardized exams encouraged teachers to make curricular decisions that would result in higher test scores for students. For example, teachers found it important to teach test taking strategies, made teaching the standards on which the exams were based a priority in lesson planning, and taught to the average student. Teachers used words like "teaching to the top of that bell curve" and "cluster management" to describe their effort to move students through tested material efficiently before the date of the summative assessments. Further, since the students are compared to other students nationwide regardless of their socioeconomic, English as a second language, or immigrant status, teachers' approach to teaching had to account for producing outcomes irrespective of their backgrounds. Teachers noted the flaws within this system. One teacher, who teaches a class which is mostly

composed of English Language Learners (ELL), stated, "They took a [standardized] test today in reading, so my scores are going to look terrible compared to somebody in a class where they're all native English speakers." While measuring reading growth is important, measuring native speakers and ELLs against the same standard likely will encourage teachers to focus on skills that will help them test better rather than speaking and listening skills that will overall better improve their English, and thus reading, skills.

Teachers reported feeling the pressure to ignore what students need and/or ignore their personal teaching approach to teach what is needed to score highly on the test that will soon occur. The urgency of covering test content and the stress of the exam outcomes impacting their professional performance led to teachers' feelings a lack of ability to change their lesson plan if the material was beyond students' zone of proximal development, not keeping the students' interests, not matching students' emotional availability, and/or not enhancing their test taking ability. Teachers also felt they often had to forgo their personal teaching approach to cover material even if it meant students would be overwhelmed by the amount of content or lack understanding about the connection between concepts and skills. As one teacher explained, "I just can't get a whole lot of stuff done because I've got to move on to the next skill. So they [can] take the test."

The emphasis on standardized testing also resulted in teachers' shifts in understanding of their own measure of student success. Only the skills that were tested by these metrics were perceived as foundational. If the subject had already been tested, teachers deemed learning in that area as complete and not necessary for further teaching. Despite the lack of consensus among teachers on whether the tests actually measured students' understanding of the standards, teachers mentioned using the test as a measure for their own professional success. For example,

one teacher explained that when scores are low, teachers would "go in and go 'okay, I must be doing something wrong.' And then they started developing these goals that eat away at their own skills and capabilities as opposed to solving some problem." Teachers viewed their own success based on the scores of students, changed their teaching strategies in response to those scores, and blamed themselves for the outcomes of those tests.

Teachers argued that qualifying and quantifying teaching and learning as measures for success was misguided. One teacher explained:

> *[State Standardized Assessments represent to me a lot of things that I think are a misguided attempt to quantify and qualify teaching. I'm not sure that I feel that they are accurate and capable of doing what the powers that be want them to do.*

Here, the teacher explained that the standardized exams represented one of many misguided attempts to measure success of teachers and teaching. Among these measurements of success included normative perceptions of successful student behavior.

Classroom Management

While understanding the shortcomings of qualitatively and quantitatively measuring student success, teachers simultaneously found themselves buying into the exact measures they wanted to reject. Due to the system itself, teachers assimilated to the notion that learning should be one aligned with the methods of learning that best match standardized outcomes. One way teachers identified buying into normative perceptions of success was by holding classrooms to the normative standards of successful student behaviors. One teacher described in a sarcastic tone:

> *They can't be too happy because if they're too happy, and having fun, if they're laughing.*
>
> *and God forbid if they're out of their seats and jumping around, something is wrong.*
>
> *They cannot be having too much fun.*

Although said facetiously, this teacher later noted that she did often internalize this judgment and apply it to her own and other classrooms. Successful student behaviors involved sitting in their seats, listening to the delivery of content silently, and being serious about learning. Teachers used phrases like, "shackled to those chairs", "formal instruction", and "not out of their seats" to describe the behavior of students in a classroom perceived as successful by administrators or outside observers, and thus, themselves.

Doing All Parts of the Job

Teachers responded that teacher success encompassed taking on myriad roles for students. This included clerical work such as submitting attendance, grading, reporting grades, completing Individualized Education Programs (IEPs) for special education students, and developing personal professional goals. One teacher explains that a stressful component is remembering all the things that you are expected to do:

> *Did you turn your attendance in the first 10 minutes, or did you forget, and have they*
>
> *called you six times? Or did you forget your lunch duty? Or have you done your grades*
>
> *weekly and turned in? And have you called the 15 children who are failing your class*
>
> *because they're not doing their work?*

Administrative work expected of teachers included calling and communicating with parents to update them on student's academic progress and behavior, overseeing students during lunch or recess, and communicating with other colleagues to ensure that students receive necessary support. This is likely due to the focus that has been placed on data-driven mandates in the

current accountability regime. Within this system, teachers are often reprimanded as not doing their job if they do not have detailed documentation of the attempts to push students toward success. If more teacher energy could be diverted to building authentic relationships and helping students connect to the material, success would be more likely to follow.

What Student Success Is, as Defined by Teachers

Doing Well on Standard, Rigidly Defined, Isolated skills

It follows that if teacher success is measured on student outcome data, that student success is defined similarly, usually by grades and standardized state assessments. Students' test scores often determined whether the student was seen as successful. Teachers recognized that this system was flawed:

> *We are judging children and teachers on a one snapshot kind of deal… if on that one particular day, a student is upset, or they aren't good test takers, or they just aren't interested, then that shows that they're failures.*

Typically, these standardized exams that are the major measure of success do not account for students' test scores over the school year, their ability to express understanding through another medium (e.g., essay, project), or their physical and/or mental wellbeing on the day of the exam. This teacher notes the limitations of this systematic measurement, indicating that this exam only looked at "one snapshot" of a student and that through this one measure, students were deemed as "failures" or successes.

The inclusion of material on these exams is usually not set by classroom teachers or even at the local level, but by state systems in place. One teacher noted:

> *It's pretty cut and dry because it's set by the state. Our benchmarks are released [standardized state] tests. And so, in terms of feedback, what students get is a score,*

right? And they know which questions they missed. And they know, theoretically, they

know which skills they struggled with.

By limiting the students to measures of success imposed by the standardized exams, the feedback students received in response to their learning was also limited. The binary of getting the question correct/incorrect or passing/failing gave the impression that students no longer needed to work on the skills they had correctly answered, and that learning could be reached in a binary of achieved/unachieved. This is a limited conception of the capacity of students. As one teacher noted, "I think the key takeaway there….The whole point of standard is assuming that everyone's going to be the same. And they're just not."

Because the normative view of student success limited students to their performance on standardized exams, learning within the classroom focused on tested content. Across elementary and secondary levels, students learned tested skills rather than skills that may be helpful in other life contexts. For instance, an elementary teacher shared:

In language arts, we're still teaching guide words. When was the last time I looked up a

word in the dictionary and had to use the guide words? Should people know a basic ABC

order? Yes, but do they really need to know which three words fit on the dictionary page?

I just think we're a little far away from reality.

Here the teacher exemplifies the tension they feel between setting a student up for success within the school system and setting them up for success in life. Since the students' success will be measured from their ability to perform on an exam, however, the content that they learn in school is also constrained to the material that will be covered on the exam. For a more in-depth discussion of the constraints of the system, see the *Facilitators* and *Barriers* sections below.

This short-sighted focus on achievement, built into the current assessment and accountability system, employs a fixed-mindset which is harmful to students because it encourages a pattern where they see themselves as failures. As one teacher explained that when focused on state standardized assessments,

> *there's very little time for them to explore and find out what they're really good at. And a lot of them have arrived at the conclusion at this young age that they're not good at anything and that really breaks my heart because they're all good at something and they don't see it because it's not valued by the educational system. My kids that can draw beautifully and are so creative, that doesn't really do them any good on their [state standardized assessments]. My kids that can take apart an engine and put it back together, that doesn't matter to anybody giving a test anywhere. And so, to them, that's not important that they can do that. And it breaks my heart because to me it's stripping them of their sense of self-worth.*

Here, students' feelings of success are an important component of their actual success, and lack of success is seen to attribute to low conceptions of self-esteem.

Reimagining Teacher Success

Interestingly, the confines of the system and the lack of power felt by teachers meant that many of them had a difficult time reimagining what success might look like in the ideal (See *Barriers)*. When given time for reflection, teachers brought up the idea that they feel success when their students are successful, though not always in traditionally defined ways. Additionally, developing authentic relationships with students, which can be theoretically described as critical care pedagogy, was a marker of success that most teachers shared.

Successful Teachers Have Successful Students

Teachers consider themselves successful when their students are successful. In the following section, reimagined pathways to student success will be explored. When teachers can facilitate student success in those areas, they themselves will be successful.

While teacher and student success are currently often defined in terms of rigid standards assessed within their subject areas, the following quotation is illustrative of what teachers across elementary and secondary levels indicated, that teachers are concerned with student development outside of their acquisition of knowledge about their specific subject: "Part of that responsibility of a teacher is to provide…educational guidance, moral guidance, to shape students' lives in a more meaningful, positive manner."

Teachers want to define their success in a multitude of ways other than those currently constrained by the system. In order to reimagine what success could look like, facilitators discussed below need to be in place so that teachers would have the freedom to be defined as successful when acting in ways aligned with their ideals.

Critical Care Pedagogy

An additional clear indicator of when teachers felt successful was the characteristic of establishing authentic student-teacher relationships. Critical care pedagogy affirms that all students, particularly students of historically marginalized backgrounds, thrive when they know they are cared for and affirmed in their socioeconomic and cultural backgrounds. Teachers' ability to make students feel this care is a marker of teacher success. A majority of teachers in this study discussed how their relationships with students not only facilitated students' academic success but were a success in and of themselves.

We see our students every day. We work with our students; we get to know our students. We have to build relationships with our students on a level that no one else will. We have to form a bond with them for them to be willing to learn from us, and so, in doing that, we form a connection with them, and we sort of take them under our wing. I joke all the time that the kids I'm going to have are my students. And that's because it's true.

In this quote, it is evident that the teacher considers not only this relationship important for his students to learn, but also significant for himself. His joking about his students being his children is indicative of the personal fulfillment and success he feels from forging those strong relationships with students. His measurement of success recognizes that his knowledge of students' lives and the resulting relationship of trust forms the foundation on which students are willing to engage in the risks and failures necessary for learning to incur. Further, his positioning as a beneficiary from the relationship with students highlights not only the deep importance of this relationship in the professional life of the teacher but also in the personal.

Teachers also recognize that their relationship with students enhances student learning outcomes. The framework of critical care pedagogy asserts that culturally additive learning communities are characterized by high-quality relationships and high academic expectations which challenges the color-evasive assumption present in White feminist ideologies of caring marked by pity and the lowering of expectations. The sense of high academic expectations, facilitated on the foundation of a strong teacher and student relationship is exemplified in the following statement:

If the kid knows that you care about them, and the kid knows that you're there for them, and that you're not treating them like a number, and that you're not treating them like a grade on a sheet. You know their name, you know what they like, you know what they

dislike, you know what they're doing outside of school. In my experience, they will, like, move heaven and earth to try and do whatever you ask them to do.

While the relationship itself is important, the teacher indicates that these relationships are a facilitator for student success and therefore teacher success. Further, the teacher recognizes that each student is unique and that to engage them, you must know the things that distinguish them from others, including what students do outside of the school setting. This recognition of difference is a key characteristic of critical care because it challenges color- and power- evasive notions of caring and generates curiosity about students' own funds of knowledge (Moll et al., 1992) rather than holding a generalized expectation that students will assimilate to the teachers' aesthetic of caring. Additionally, the teacher expresses interest in the students' knowledge outside of the school, revealing that relationship building with students expands to who the students are beyond their student identity. The affirmation of students' identities and interest in their contexts define the critically caring relationship between the student and teacher and act as the building block for students' academic success.

Reimagining Student Success

When teachers discussed how they wished to define success for students, they talked about how they wanted students to have agency to learn about things that they were interested in and to focus on growth and their academic and/or vocational areas of strength. In doing this, students could foster their own feelings of success, which are tied to feelings of self-worth.

Feelings of Success

One piece of accountability that is often missing from school definitions of student success is how successful students feel. In order for this idea to be incorporated, the system needs to not only depend on external assessments but allow space for internal inventories for

students to reflect upon and measure their own successes. By adding this component to systems of accountability, states could show evidence that they value student voice and understand that intrinsic motivation is evidenced to be a key indicator of future success. As one teacher explains:

> *We want them to take responsibility for their learning to become adults of responsibility. A part of that is learning how to look honestly at your own work, at your own performance, your own yardstick. Where are you; how are you growing? ...Whether or not you're giving 100%. You know where you have areas to improve, and you also know where you do well. So, I think that we need that type of assessment as well, not just an outset, not just an extrinsic evaluator and extrinsic assessment, but an internal one too.*

This is not without measurement or accountability, but the accountability looks different than the rigid system in place today. She elaborates:

> *Measuring anybody's knowledge and gifts and skills on a single day, in a single way, at a single moment, against a single standard, with a single rubric— that's [she makes a tsk noise] that's incredibly reductionary... It diminishes us tremendously.*

Many teachers confirmed that they believe in continued measures of success through tests and projects, but that one summative test built around rigid standards does not allow for students to show the ways in which they are successful. When tests can be a piece of gathering information about success, but not seen as the sole indicator, it will set students up for greater success.

Agency and Exploration

Teachers repeatedly discussed that student success will not look the same for every student. This was often discussed in opposition to the rigid system teachers are educating within. In order for students to be able to explore what's interesting to them, they first have to be exposed to those ideas. As one teacher explained, "One of the ultimate goals [for students] is

college readiness and college, but I don't think that college is for everyone, so I think that exposing students to different ideas for future paths is super important." If teachers are to provide students with a variety of ways to be successful, they also need to provide them with the agency to explore and learn about ideas that they are interested in. To illustrate this point, teachers said things like:

> *You'd have to get to know your students first to see what they're interested in. Let's say, for example, this year we have a student who's really interested in lawn care and excavating and that sort of thing. Or if kids are interested in, say, video games. Okay, you could find some articles about video games, but it could be more about how they're designed or how they're created. So, there'll be more of a process as opposed to just, well, here's a nonfiction article about why birds fly to the south. Well, who cares? We need more interest.*

As is true within systems thinking, we see how this pedagogical practice intersects with critical care pedagogy as a pathway to student success. When teachers value their relationships with students, then students understand that their teachers want them to explore their own interests and be successful and are more able to do so. The contradiction to the normative system is also evidenced here. While student interest can be incorporated to some degree in the normative system, the freedom to deeply explore those interests is limited. Another teacher noted that she wants to "lead the students into their own learning, to get them to want to buy in and learn." This is more strongly aligned with systems that focus on competencies rather than explicit standards that have to be taught to each student in a particular way.

Even more important than maintaining interest, though, was how presenting information that is not promoting student success is impacting students' self-conception. Teachers spoke

about how the lack of student agency along with strict standards can harm self-esteem because constant failing and need for improvement is detrimental to developing self-worth. In an ideal situation, one teacher shared she would prefer:

> *a much more student directed sort of a layout. I think that there will be a lot more room for kids to explore different types of vocations, occupations, hobbies, interests in order to kind of find their niche because there's a lot of students that I know don't have that opportunity. They don't have it in school right now because we're so busy force feeding them what they have to know, and they don't have it at home because they don't have that support from whoever they live with.*
>
> *And so, I think that that's probably something that they would have to start when they were really young, is, you know, empowering them. To say this is what I'm good at, and this is what I'd like to do, and this is what I'd like to do with that, and this is what I'd like to try, and teaching them that trying things is okay and deciding that you don't like them is okay, too.*

This agency to explore would lead to students who were able to grow specifically in their strengths, which is related to the next theme within reimagining student success.

Growth and Strengths-Focused

In a reimagined system, students may have the freedom to not spend the majority of their school lives constantly remediating their weaknesses, but instead realizing their strengths and allowing for continued growth in those areas. Here, assessment could facilitate decision making and growth not just for the teachers but students, too, as opposed to just a summary of learning. There is empirical evidence to support that focusing on strengths leads to greater rates of growth

than remedial education.[2] One teacher spoke of the contradiction of this ideal to the normative system, *"I like to see something that builds a student's understanding of what they're going to do or what they could do as opposed to something that says this is what you learn. Period."* Another teacher ties this growth mindset to assessment and accountability:

> *What if we… measure growth? Then in a couple of things happen. Number one, our students become less risk averse. Our students right now are very afraid to go out on limbs because limbs break, and the way that we measure students now doesn't allow that. We have a pass/ fail mindset now, and the emphasis is on passing. You get no points for growth, you get no points for coming close, you get no points for being creative, you get no points for being unusual. You know, it's either you pass or you don't. So if we had a growth mindset, where we could see how our students perform on these standards at a given level on a given day. Then we identify areas where that if they're making the mark and we challenge them to surpass that, we challenge them to, because those become their strength. Those become their gifts. A lot of our kids don't know their gifts.*

This daily measurement is not as systematic as the assessment and accountability system of today, but it allows for much *more* data. While data-driven decisions have often been tied to benchmarks and standardized data, this reimagined system of assessment and accountability allows the teacher to capture much more holistically what students know, and then encourage

---

[2] Hodges and Harter (2005) provide evidence that growth happens in greater magnitude when students focus on growing skills associated with areas they excel in rather than focusing solely on remediating their weaknesses. A comparative dissertation study in the mid-1950s about three reading improvement strategies stumbled across this finding unintentionally (Glock, 1955 as cited in Clifton & Harter, 2003). While modest gains were made in reading speed and fluency by those with low-level ability during the pre-test, huge gains were made by those who already had an aptitude for reading quickly. These findings were applied to the market, which led to a push for managers to use strength-based approaches with their employees, leading to increased productivity and less turnover.

students to grow and use those successes in ways more aligned with life outside of school and in the real world.

Authentic Problem Solving

Several teachers noted that the education system would benefit from a shift to focus on fixing authentic problems, saying things like:

> *Let's give them real life problems. Let's give them authentic consequences and authentic audiences. And let's see. Let's open up opportunities for them to show us things that maybe we didn't even think to assess in the first place.*

Many teachers made mention of the fact that "real life success" is more connected to skills other than those assessed on standardized assessments. When discussing critical thinking, creative thinking, communication, collaboration, and citizenship skills, one teacher noted:

> *That's what's important to me…Those are the things that are going to translate into our real life success and are really going to make or break [students'] ability to perform in any type of environment.*

Teachers often joke about how life's issues don't come with multiple choice options. This theme indicates that teachers want to set students up for success by presenting practice for solving problems that more closely mimics what they will encounter outside of school. This idea further confirms the first theme, because when students can ask authentic questions and problem solve to find answers, it will assist them in feeling more successful. In turn, this will support the reimagined version of teacher success, which is defined by helping students to find success in these ways.

Discussion

Throughout the results, it is evident that the certain system of assessment and accountability constrains teachers' definitions and aspirations of success. Dialogic experience with that system leads teachers into a new imaginary for success. However, the impact of the new imaginary is negatively impacted by the current accountability system and thus needs to be aligned.

Facilitators for Success

Teachers and students alone cannot change the system to redefine success. There are components of the system that need to be amended in order to support these reimaginings. CSH encourages an investigation of who holds power and who is advocating for those who are without. These facilitators would support the reimagined versions of success by moving toward a common vision across system stakeholders.

Supportive School Staffing

Administration. One of the most important staffing decisions in order to facilitate this new reimagined success for teachers and students is the administration. This is true at the building, district, state, and national level, all of which impact this reimagined state. Administrators who support teachers and allow them the freedom to try new things will create an environment with less fear and more willingness to change. Policymakers who build flexibility into the system will encourage administrators who are supportive of teachers in this way. Alternatively, policymakers and administrators who do not take this approach become a barrier to implementation. Many teachers noted that unsupportive administration was a major part of teachers experiencing burnout during the pandemic. Key characteristics of a supportive

administration include allowing for flexibility, encouraging an environment where it is safe to try and explore new ways of teaching, and helping teachers define success for themselves.

Teacher Support. An interesting staffing role mentioned in interviews included adding support staff not only for students but that teachers could use as well. Teachers, who were interviewed during the pandemic, expressed the role of burn-out in teachers' choices to leave the classroom. This burn-out was created by the high-stress, high-level expectations with minimal support. A way of helping support teachers would be to allow access to mental health support during the school day where teachers would be able to discuss issues related to their classrooms with staff who were able to understand their context and offer support in ways that are feasible within their school environment.

Humanizing Teachers and Treating Them as Professionals

Teachers are more likely to find success when they are treated as professionals who know what students' needs are and how to meet them. Again, this requires teachers to be given the flexibility and freedom to experiment. Additionally, this requires fiscal changes at the state-level. When less money is spent on testing contracts, more money is available for attracting, paying, and keeping a professional teacher workforce. It is important to remember that changing ideas and exploring new ways of teaching will not always be successful in the first iteration.

Barriers to Success

The largest barrier to leading teachers and students to success is the system itself. There are a number of ways that the current system and all of the actors within it perpetuate the current definitions of success that are held. These barriers will need to be addressed via the supportive staffing models detailed above in order to move towards the reimagined definitions of success.

The System as Barrier

One of the biggest barriers to success is the fear of failure, and that failure is determined by the system as is, especially the systems of standards, assessment, and accountability. Teachers are often uncomfortable letting go of ways of teaching that have "worked" in the past, even if a new method might allow for higher levels of success for their students. This is especially true when salaries, recognition, and requirements are structured around outdated notions of success related to standardized assessment. As an example, one teacher shared:

> *I think I'm just as guilty of it as any lawmaker or parent or the community that I*
>
> *complain about not supporting. I'm part of that same thing. I want to know if my kid has*
>
> *done well. "Is my kid an A student?"*
>
> *And the teacher going, 'Man! Your kid really got so much better at writing!'*
>
> *'Well, did they pass?!' And that doesn't necessarily need to be the question. The question*
>
> *is, does my kid still love reading enough to come back and want to do it next year? And*
>
> *that's not the question I'm going to ask as a parent. I'm going to ask, 'Did my kid pass?'*

The system itself constrains what is prioritized when considering success because of the measurements of assessments and the accountability tied to them. Teachers shared that even when they believed certain skills (e.g., writing) were more important than others (e.g., guide words), they prioritized teaching the skills they knew would be on the test even if they weren't particularly helpful to the students' overall knowledge and development. One teacher even noted how this aspect of the system de-incentivizes doing the *right* thing for actual student success over success that is measured:

Teachers [that] are usually a little more holistic, I think that's where they get in trouble.

The growth that they create in the students may not be as valuable on the test. But it's

what the kid needed at the time, and so that gets to be more difficult.

A system where teachers are getting "in trouble" for giving students what they need at the time is a major indication of the need for critical self-reflection and change.

Success for Who? An additional barrier to moving away from standardized, rigid definitions of success is that it opens up space for disagreement. When the belief that every individual is unique is central, it may be difficult to have consensus on what success looks like by every teacher or every student. Parents and teachers are hesitant to break away from a system especially when concerns around evaluation are still tied to the "old" system, as illustrated by the teacher above. With new flexibility, administration needs to be prepared for disagreement over what success is and have pathways in place for helping to assuage tensions around differing definitions of success. There are systems that are successfully navigating these tensions, such as networks of project-based learning schools that utilize rubrics and professional development communities. Additionally, it is necessary to consider how these more subjective systems may allow opportunities for racism and classism. Safeguards need to be built into systems in explicit ways to ensure that the outcomes of the system are not advantaging and disadvantaging students on the basis of race or economic status.

"Out of Touch" with the Ideal

When asked about what an ideal education system might look like, one teacher laughed, exclaiming, "You know, I'm so out of touch with what ideal might look like!" Several other teachers had similar remarks when asked about reimaging success. This is evidence that the system is so powerful, that many within it are not able to think or dream about what could be

better. Rather, they just put their heads down and do what is required. One teacher noted, "I think we're so mired down in 'this is how it's always been done.'"

This often leads to viewing students with a deficit mindset. Even though through interview data it is obvious that teachers by-and-large want what is best for their students, they would then report things such as "they're never going to pass the fifth grade [standardized] test."

It is here that freedom dreaming (Kelley, 2022; Spaulding et al., 2009) is desperately needed. All members within the system need to, as supported through CSH, try to look at the larger system, understand it from perspectives other than their own, and dream of how the system can be reimagined to work better for all of those within it.

We know that teacher success is confined by the limitations of the existing system. While teachers have agency in their classrooms, to build relationships with their students and to make decisions on what and how to implement lesson plans and classroom systems, they are unable to engage meaningfully with the systems that drive problems that are beyond teachers' control. For example, in communities where rates of community violence are high, students are at increased risk of exposure to trauma and its related symptoms (e.g., avoidance, numbness, irritability). Students who express these symptoms in the classroom or are triggered within the school day, are more likely to struggle to meet the normative standards for success as defined by standardized assessments. Teachers are also likely to struggle to support the students with adverse experiences in supporting standardized measures of success, as the source of their trauma stems from outside of the classroom. Teachers are thus asked to figure out ways to survive, exist, and/or succeed within the confines of these parameters. Furthermore, schools are also limited to measures of success that are within the confines of the existing system. Schools are not designed to support students who have needs that are non-academic, because success is

not measured with non-academic standards. It is here that critical care pedagogy becomes a key component, to envision caring for students as part of the necessary structure of the classroom.

Teacher Time is Valuable and Limited

It is important to acknowledge that this reimagined way of defining student success, where there is not one clear pathway, requires more planning time and a different kind of teacher stress and burden. Even in the school system as it is, all the requirements of teachers outside of teaching can be a barrier to success (e.g., meeting students' non-academic needs, staff shortages requiring additional coverage time, etc.). Stress and pressure to be successful impacts actual success, and administration will need to develop clear communication systems in order to not continue to burden teachers with more expectations without additional measures of support.

Conclusion

While the normative perceptions of teacher and student success are largely rooted in rigid uses of standardized measures, teachers and students are yearning to function within a system where they can bring their whole selves. In utilizing CSH, and considering not only how the system is, but how it ought to be, success can be reimagined. In this reimagined system, student success is measured both externally and internally, providing students the agency to explore their interests and focus on growth, especially in areas where they already have innate abilities. This approach respects students as humans with the capacity for their own decision-making. When teachers can promote these kinds of student success through critical care pedagogy, this can be a means of their own success. In order to facilitate this, systemic change is needed. Supportive school staffing is key. Additionally, teachers need to be seen as whole people who bring their personal and professional selves into the teaching profession and can be trusted to make decisions that propel student success. As one teacher shared about teaching in the pandemic:

School can, learning can take a lot of different forms. It could go a lot of different

directions; it can use a lot of different forms. And so, if it can be those things, why on

earth wouldn't we let it?....We've broken the mold. The question is now what?

The "now what" supported by this research is to humanize teachers and students and redefine

success, so that individuals are not just a cog in a machine but a moving individual within a

beautiful, diverse landscape. This post-pandemic system of assessment and accountability may

not be easily measured or tracked, but when we allow individuals room for agency and growth,

the collaboration and new creation of this reimagined system is one in which creative solutions

to the *mess* of global problems, inside and outside of education, can better be addressed.

References

Ackoff, R. L. (1974). *Redesigning the future: A systems approach to societal problems.* John

Wiley & Sons.

Antrop-González, R. & De Jesús, A. (2006). Toward a theory of critical care in urban small

school reform: examining structures and pedagogies of caring in two Latino community-

based schools, *International Journal of Qualitative Studies in Education, 19*(4), 409-433.

https://doi.org/10.1080/09518390600773148

Clifton, D. O., & Harter, J. K. (2003). Investing in strengths. *Positive organizational*

scholarship: Foundations of a new discipline, 111-121.

Gallup (2014). State of American schools. Retrieved from http://www.gallup.com/

services/178709/state-america-schools-report.aspx.

Greenberg, M. T., Brown, J. L., & Abenavoli, R. M. (2016). Teacher stress and health effects on

teachers, students, and schools. Edna Bennett Pierce Prevention Research Center,

Pennsylvania State University, 1-12.

Guggenheim, D., Kimball, B., Chilcott, L., Strickland, B., Canada, G., Rhee, M., Weingarten, R.,

Roland, E., Richman, B., Finton, G., Cassidy, J., Roberts, K., Beck, C., & Legend, J.

(2011). *Waiting for "Superman*." Paramount Home Entertainment.

Hodges, T. D., & Harter, J. K. (2005). The quest for strengths: A review of the theory and

research underlying the StrengthsQuest program for students. *Educational Horizons*, 190-

201.

Instructure. (2022). The state of teaching & learning in K-12 education. Instructure.

Jennings, P. A. (2020). *Teacher burnout turnaround: Strategies for empowered educators*. WW

Norton & Company.

Johnson, C. S. (2014). A critical care ethic. In P.W. Orelus & R. Brock (Eds.), *Interrogating Critical Pedagogy* (pp. 127-149). Routledge.

Katz, S. (1999) 'Teaching in tensions:' Latino immigrant youth, their teachers, and the structures of schooling. *Teachers College Record, 100*(4), 809–840.

Kay, K., & Boss, S. (2021). *Redefining student success: Building a new vision to transform leading, teaching, and learning*. Corwin Press.

Lieberman, M. (2021). How Staff Shortages Are Crushing Schools. *Education Week*. Retrieved from https://www.edweek.org/leadership/how-staff-shortages-are-crushing-schools/2021/10

Kelley, R. D. (2022). *Freedom dreams: The black radical imagination.* Beacon Press.

Killam, L. A., & Camargo-Plazas, P. (2022). Revisioning assessment and evaluation in nursing education through critical caring pedagogy: using authentic examinations to promote critical consciousness. *Advances in Nursing Science*, *45*(1), E15-E30.

Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom*. Beacon Press.

McKamey, C. (2004) Competing theories of care in education: A critical review and analysis of the literature. American Educational Research Association Conference. San Diego, California.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice*, 31(2), 132-141.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. *The Elementary School Journal*, 84(2), 113-130.

Oberle, E., & Schonert-Reichl, K. A. (2016). Stress contagion in the classroom? The link between classroom teacher burnout and morning cortisol in elementary school students. *Social Science & Medicine, 159*, 30-37.

Reimers, F. M., & Chung, C. K. (Eds.). (2019). *Teaching and learning for the twenty-first century: Educational goals, policies, and curricula from six nations*. Harvard Education Press.

Reynolds, M., & Holwell, S. (2010). *Introducing systems approaches. Systems approaches to managing change: A practical guide*.

Rolón-Dow, R. (2005). Critical care: A color (full) analysis of care narratives in the schooling experiences of Puerto Rican girls. American educational research journal, 42(1), 77-111.

Spaulding, E. C., Adams, J., Dunn, D. C., & Love, B. L. (2021). Freedom dreaming antiracist pedagogy dreams. *Language Arts, 99*(1), 8-18.

Steiner, E. D., & Woo, A. (2021). Job-related stress threatens the teacher supply: Key findings from the 2021 State of the US Teacher Survey. RAND.

Tamayo, P. D. (2016). Report to the legislature: Digital citizenship recommendations. Office of Superintendent of Public Instruction. Washington.

Thompson, A. (1998) Not the color purple: Black feminist lessons for educational caring. *Harvard Educational Review, 68*(4), 522–554.

Ulrich, W. (1996). A primer to Critical Systems Heuristics for action researchers. *Independent Research Monograph*, Centre for Systems Studies, University of Hull.

Ulrich, W., & Reynolds, M. (2010). Critical systems heuristics. In *Systems approaches to managing change: A practical guide* (pp. 243-292). Springer.

Valenzuela, A. (1999) *Subtractive schooling: U.S.-Mexican youth and the politics of caring.* SUNY Press.

Wagner, T. (2010). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need-and what we can do about it*. Basic Books.

Zamarro, G., Camp, A., Fuchsman, D., McGee, J.B., (2021) How the pandemic has changed teachers' commitment to remaining in the classroom. Brookings. Retrieved from https://www.brookings.edu/blog/brown-center-chalkboard/2021/09/08/how-the-pandemic-has-changed-teachers-commitment-to-remaining-in-the-classroom/

Zhao, S. (2012). Global competency leadership and an empirical study of Chinese managers. In *China's reality and global vision: Management research and development in China* (pp. 104-124).

Epilogue

Standards, assessment, and accountability policy have come a long way since the 1980s excellence movement. We have seen waves and ebbs and flows from the rise of high stakes assessment and accountability policies in the 1980s through the 2000s. Over the last ten years, there has been some ramping down from such high stakes accountability systems. We have seen some expansion in the development of lower stakes accountability and performance assessments systems. In addition, we are now seeing challenges based on the high costs of accountability systems in which all students are assessed every year. With school correlations between years of testing around .90, there is little evidence to support the need to test annually for accountability (Gibbs et al., 2023), especially when it is so expensive, and the focus could be more beneficial if directed toward student growth and development. We seem to be at a crossroads nationally. We cannot ignore the pressures of neoliberal capitalism and its influence on a global culture of competition translated though national and state assessment policies.

The conclusion of this dissertation study leaves us with a picture of Virginia as stuck in the in between world of traditional content-based, standardized assessment system and the world of 21st century learning standards like the Profile of a Virginia Graduate, the 5 C's, and the necessary revamping of its accountability and assessment system to include authentic and performance assessment. This is a political space with implications for the goals of education and the future of Virginia.

Recommendations

Based on the findings of the three papers as a whole, there are a number of recommendations in response to Flyvbjerg's (2001) question, "what *should* we do?" (pp. 136-7). These include the need for greater teacher and student agency, policies that promote more

control at the local level, and a culture that promotes learning and growth rather than fear of failure.

Teacher and Student Agency

Especially in the third paper, teachers indicated their desire to measure their own success, not by external standards, but by caring for their students and guiding them to their own definitions of success. With more professionalization of the teacher workforce, teachers, with their long-term observations of students each day, should be trusted to know the relevant data of their students' learning. While the policy climate constrains teachers to act as passive receivers of policy, decentralization of policy control could allow for more agency at the individual level, both in determining the means and the assessed ends of the given standards.

More Decentralized, Local Control

Darling-Hammond and colleagues (2014) note that the ideal accountability system would hold more far-removed policy actors accountable for funding the necessary work of local, professional teachers measuring authentic student learning. In this system, teacher and student agency are trusted, and those schools where there is more work to be done are given more funds in order to best support the learning of students. Despite top-down control for the last thirty years, schools' current measures of success remain fairly consistent from year to year (Gibbs et al., 2023) and state takeover typically has either no or negative impact on student achievement (Schueler & Bleiberg, 2021). With this in mind, it is time to consider a new system of control where localities are trusted to know the needs of their community and given the resources to meet those needs.

Culture of Growth

These recommendations will require a shift in political culture. As seen in paper three, teachers have been so embedded within the system that it is hard to imagine it any other way. To encourage any change to happen, intentional work to shift culture will be necessary. Kezar (2011) lays a framework for how to implement organizational change via grassroots efforts that includes all members of the community. Additionally, programming for increased information about assessment that can help with these shifts is available (Virginia Association for Supervision and Curriculum Development, 2022a). This type of intervention will be necessary if the culture is really to change. From the CPA conducted in this dissertation, it is important to remember that power and politics will need to be attended to in these considerations of policy change. Key strategies to listen to teachers and families, and not just push top-down transformation will be needed if transformational effects are to have lasting impact. In many fields, such as engineering, failure has been embraced as part of the learning process. It is time for education to embrace this understanding, that we must fail in order to reach the best versions of ourselves and our society, and that punitive connections to failure is antithetic to the growth and love of learning educators claim to want for their students.

<div align="center">

Contextual Factors

</div>

In considering these changes and how they may come about, it is important to remember the culture and context that led to the development of the system as-is. Only when understanding how the policies were formulated can they be appropriately dismantled and the root of the issues addressed.

Attention to the Virginia Context

The Virginia context and traditionalistic political culture (Marshall et al., 1989) cannot be ignored when thinking about change in assessment policy. With historical origins in colonial plantations, Virginia must reckon with its culture of elitism that has led to top-down management in government and industry and its persistent disdain of the role of the ordinary citizen in the policy process. This elitism is baked into Virginia's political institutional structure as evidenced in such procedures as the political appointment of members of the Board of Education and Superintendent of Public Instruction in the Department of Education. While some states elect these positions, vestiges of structural elitism remain in the State's political culture. This culture will make the aforementioned policy changes even more difficult, and grassroots coalitions will likely be necessary for any type of lasting change. The development of public unions may influence these policy processes.

Persistence of Racism

If Virginia institutions, and the educational policy system that emanates from those institutions, were founded in a culture of white supremacy, and continue to be influenced by such racism, the connection between institutional racism and the standards, assessment, and accountability policy system requires further examination. In order for any of the above recommendations to have the needed impacts on Virginia's students, the historical and persistent racism within the state needs to be addressed. For instance, in the spectacle of the History & Social Studies standards rewrite, students have indicated a wish to learn both the good and the harmful history of Virginia in real ways, and at the policy level and at the school-level, individuals need to be able to name and recognize the ways that racism has long-shaped the system of education. Without the willingness to address these issues in more than symbolic

ways, change will not happen. Those resisting the system will need to look for opportunities of interest convergence to push for change with a mind to safeguard against backsliding when interests may no longer converge. While change in practice is unlikely to be swift and perfect, there is no opportunity for progress without persistent pursuit.

Neoliberalism

In paper one, I discussed that since the 1980s, there has been an increased national effort to move public schools toward privatization, manufacturing a public education crisis to push for voucher programs and defund public schools (Berliner & Biddle, 1997). This crisis manufacturing was intentional. Neoliberalism manifested in U.S. education policy in general, and in the national-wide standards, assessment, and accountability movement. As Friedman (1982), who proposed school vouchers for choice, stated, "Only a crisis – actual or perceived – produces real change" (p. ix) While this economic push was seen as a key component, the push for privatization in education had a more sinister genesis, with mandated desegregation and the massive resistance that followed. MacLean (2021) traces the connections of the school voucher system to resistance to integration starting in the mid-1950s, with a special emphasis on that connection in Virginia. As she states, Friedman and other proponents of school choice:

> saw in the backlash to the desegregation decree an opportunity they might leverage to advance their goal of privatizing government services and resources. Whatever their personal beliefs about race and racism, they helped Jim Crow survive by providing ostensibly race-neutral arguments for tax subsidies to the private schools sought by white supremacists. Indeed, to achieve court-proof vouchers, leading defenders of segregation learned that the best strategy was to abandon overtly racist rationales and embrace both an anti-government stance and a positive rubric of liberty, competition, and market

choice. The pattern took shape most clearly in Virginia, the driver of a regionwide

strategy of "massive resistance" to the mandate to desegregate public schooling.

(MacLean, 2021, p. 4 -5)

This provides additional evidence of the harm of race-neutral policies, and why centering

conversations about race in education policy is so important. At the time of this writing, we see,

again, resistance to discussions of equity and a drive toward privatization in Virginia schools. In

a discussion with Senator Hashmi of the Virginia Senate in Spring 2023, she confirms this is true

today:

> If the goal, and we've seen this from our Republican colleague and from the
>
> administration, if the goal is to move public education away from the sphere of public
>
> education, and more towards privatization, then the goal of assessment is to demonstrate
>
> the failures of the system, and that is fundamentally at odds with many, many people, and
>
> I'm speaking on both sides of the aisle, who want to use assessment in the ways that they
>
> are designed to be: to give us opportunities to see where the gaps are, to address those
>
> gaps, to identify where resources are quite necessary, where we're doing well …. Once
>
> you begin to politicize it, and you're using assessments to build a narrative that moves the
>
> needle in one direction or the other, then that's not the goal of assessments, and it's a very
>
> unfair proposition to put on our education system and on our students as well. (personal
>
> communication, April 24, 2023).

The results of this dissertation have provided evidence for what happens when degenerative

politics enter the foray of education policy making, and the biggest loser of the policy process,

Virginia's public school students, becomes the very purported beneficiary of the system. As

policymakers reflect upon the duty that has been placed upon them to serve students, a re-

consideration of their role and a willingness to disengage from the politics and center students is unlikely to happen outright but would be ideal. However, there are particular actors and coalitions that can still help push for change.

<h3 style="text-align:center">Implications for Practice</h3>

There are several recommendations to consider as next steps for assessment activists and advocates for Virginia's students. With the creation of the Profile of a Virginia Graduate and movements within the assessment and accountability system related to developing students who have the skills described in the Profile, there is a window of policy opportunity. In order to seize this opportunity, there are three key tasks that follow: consideration of what that system might look like, coalition building to push for that system, and waivers from the current accreditation model to allow the freedom for implementation to fidelity for school divisions who implement such a system.

The recommendations based on this research have been supported by some policy stakeholders' suggestions about the direction of schools in Virginia. As an example, the Virginia Association for Supervision and Curriculum Development (VASCD) made its top three priorities for policy advocacy in 2023 to promote autonomy for teachers, modernize the assessment system, and personalize professional development for teachers (VASCD, 2022b). Within their priority of modernizing assessment, they referenced a summit that occurred over the summer of 2022 in which stakeholders, educators, and assessment experts from Virginia gathered in order to discuss directions for Virginia's assessment system after the enactment of HB 585 (VASCD, 2022a). This is a hopeful sign that advocates for assessment change will be active members of the policy process. The hope was this work could inform the Virginia HB 585 work group as they make their assessment decisions. Their main recommendations included ensuring that the

assessment system captures "proficiency, growth, and performance" which includes helping all stakeholders involved to better understand the purposes of assessment (VASCD, 2022a, p. 3). Additionally, they noted that "performance-based assessments must become a cornerstone for measuring student learning outcomes" and that the incorporation of the 5 C's into classroom pedagogy and assessment was a must (VASCD, 2022a, p. 5). Within this discussion, there was also an acknowledgment of the negative impacts of the current accountability system and a call to reframe assessment to be used for decision-making rather than punishment. In order for this system to work, they acknowledge that teachers must be trusted to make informed reports about the proficiency of their students.

A Virginia High School Full of Ideal Virginia Graduates

bell hooks (1994) notes that the school she first attended was full of teachers who knew her and her family well and taught her in the context of that knowledge; she says, "attending school then was sheer joy" (p.5). Sheer joy did not mean the absence of conflict, but when she felt seen in school, and was allowed agency for exploration, there was a love of learning. Many students today do not feel this joy, as noted by teachers in paper three who spoke to the way that the system as-is strips students of their self-confidence. So, what would it look like to allow students that space to find joy in learning? The following vignette is a consideration of what a high school might look like if fully aligned with the Standards of Learning unmoored from the SOL tests and instead redirected toward the Profile of a Virginia Graduate. This vignette puts into imagined practice the system as recommended by teachers in paper three.

*When you walk into Ideal Virginia High School (IVHS), the first thing you notice is that it is **not** quiet. The buzz of the classrooms can be heard in the hallway, and it is clear that none of the rooms have students silently sitting in their seats receiving the day's lessons. After checking-*

in, you peak your head into a science classroom. Students are gathered around lab equipment and are discussing collaboratively the next steps they want to take on their experiments. You notice that the materials within the different groups are not the same, and as you walk around the room, it is clear the different groups are engaged in different experimental set-ups. The teacher comes over and explains that within this chemistry unit, students developed their own questions about how and why materials interacted. After submitting the lists to the teacher, they were provided their materials and were currently testing different hypotheses they had co-constructed. One group begins a rather heated argument over what they think should happen next. The teacher walks over and helps to de-escalate the situation, practicing compromise and determining how to move to the next step when group members are not in agreement.

Next, you head across the hall into a history classroom. Here, students are in a class discussion about a recent Supreme Court Case. You look around nervously when one student mentions access to birth control and cites historical changes that have happened that may have informed the recent decision. Students did not talk about these topics when you were in high school, at least not in the classroom! Another student brings up a different perspective that informs her decision. The teacher acknowledges the differences in opinion and encourages both students to think critically about their own perspectives and those on the other side. What might be informing these decisions? What other examples have they talked about this year in which a similar situation had happened at the Supreme Court? Students all turn and brainstorm with those around them. You hear students bringing up court cases you have never heard of, but you also notice they are much more engaged with and informed about the topic than many of your peers. You're interested in how students learn about communicating here at IVHS.

This leads you to an English classroom, where students are discussing poetry. Ah, poetry, you find yourself thinking. I know what's coming here. Similes and metaphors and selecting from a list of choices about what the poem could be about. Again, the classroom experience is unlike what you're expecting. Students are working in partners and are reading and rereading a sonnet by Shakespeare. The assignment? Students are to follow the figurative language patterns in the sonnet but re-write it about a topic that is of interest to them and their partner. As you walk around, you hear students identifying the figurative language and flexing their creativity as they write about everything from the latest comic book-based film to their stress over social media. The partner work allows for each student to bring their own talent, whether that's in the creative, the concrete, or somewhere in between.

In thinking of the concrete, you head to an algebra classroom. Sure, the humanities can utilize this way of learning, but what about math? Math is just the facts, right? As you walk-in, you notice students holding measuring tape. As you walk over to a student and ask what's going on, they inform you that they have been tasked with determining the materials needed to build the environmental classes' new garden. They must determine the amount of lumber and soil to build and fill the beds, and then work-up a budget that will be submitted to the school in order to purchase the necessary materials. Talk about authentic learning! You think back to your years of finding area and perimeter. You wonder, what will happen when there are no more school projects? So, you ask a teacher, who notifies you that at the beginning of the year, they also ask for projects from parents and the community. While not every single skill is attached to a community project, there is time in each class for working on learning new skills and for working on community projects.

As you leave that room, you find yourself thinking, this is all well and good for one day, but what about ensuring that they know the content? How will they be successful if they don't know all the basics. Then, you find yourself reflecting on what you've seen. How students in science were applying their skills to solve problems that interested them. How students in civics were discussing current events and drawing from history as evidence of their points of view. How students in English seemed much more interested in metaphors and their application than any high schooler you remember reading a Shakespearean sonnet. How in math, students were using algebra, not to determine something for Mr. Abstract's farm in a textbook, but for the school garden to be placed outdoors and used by the environmental science class. Here, you think, lies maybe the more important set of skills. You see how the connection and the reality could allow students to hold onto their meaningful learning in ways you never experienced in your own high school life, and how maybe, just maybe, this new way of schooling might be worth a shot.

Waivers

So, how to make this happen? Schools can't operate both within the system as-is and create a new system. So, waivers are necessary in order to allow the freedom to try and explore new ways of schooling. Virginia is allowing waivers for some programs, and a key component of these waiver applications needs to be exemption from SOL tests. The New York Performance Standards Consortium, in operation since the 1990s, has seen success with this model, where students only take one English standardized test and through waivers are exempt from all additional standardized testing. A network of schools is implementing project-based learning across the state, but these classes are not yet freed from the constraints of the accountability system. In order to gather more information about this way of schooling, this freedom must be

prioritized. Additionally, the political culture in Virginia differs from that in New York, so as noted above, special considerations about how to advance such policy innovations within the culture of Virginia's policy system will need to be addressed.

Necessary Future Research

In order for these recommendations for assessment policy innovation to succeed they need space to be implemented and piloted. Future research should consider the needs of localities. For example, Kentucky is currently piloting flexible local control in some districts, and research to learn from these implementations and such work is highly relevant to the Virginia context as they share some elements of political culture. Additionally, learning from systems like the New York Performance Standards Consortium, which has been operating with waivers under more local control since the 1990s, would allow for understanding the long-term outcomes of such policies, as well as identifying barriers and facilitators to success for the Virginia context.

If such waivers are granted to Virginia schools, further research could be conducted in pilot schools. Information gathered in these settings would best inform to how this policy transformation might look in the Virginia context. Additionally, this dissertation supports the notion that determinations of how local policy actors, especially division leaders, see their roles in policy implementation could illuminate where there are spaces for addressing the politics of the policy system. Since papers two and three discuss the teacher-level implementation and understanding, district and building perspectives would improve the understanding of those in the middle of policy-in-action and additional creators of policy. A next step in the research should include interviews and surveys with school board members, school division superintendents, and principals as well as other stakeholders such as the Virginia Association of

School Superintendents, Virginia Association of School Boards, and Virginia Association of

School Principals, as well as other key stakeholders in the Commonwealth.

Additionally, as discussed in paper three, when systems of assessment and accountability

are less standardized, there is the potential for discrimination by practitioners. We know,

however, that the system as-is is systematically privileging certain students at the expense of

others. It will be necessary as new systems are piloted to turn from previous color-evasive

tendencies and implement safeguards to protect historically marginalized students. This will

include looking at inputs and outputs of the education system, conducting race-centered policy

analyses as advocated for by Bradbury (2020), and considering the system at large through the

use of Critical Systems Heuristics (Ulrich, 1996) applied to the U.S. system of education as

detailed in Table 2 within the Linking Document.

Conclusion

The assessment policy system envisioned here is idealized and not yet in existence.

Virginia is struggling with its future direction. However, Love (2019) reminds us that "antiracist

pedagogy combined with grassroots organizing can prepare students and their families to

demand the impossible in the fight for eradicating these persistent and structural barriers" (p.19).

Demanding an educational system in which students can again fall in love with learning is a lot

of work, and may even seem impossible, but it is a must for our future. This dissertation has

highlighted the issues of degenerative politics and assessment policy. However, it has also

highlighted the introduction and possibility of more expansive visions of standards for engaged

21st century learning, as well as the potential for progressive policy changes with regards to the

Profile of a Virginia Graduate, the 5 C's, and the connected performance assessments. If we can,

through successful advocacy, free teachers and students from the shackles of high stakes

accountability and testing, we can provide space for agency and growth in which schooling may become a humanizing place, where students learn about themselves and the world around them, rather than a techno-rational mastery of disassociated facts leading to a lack of imagination and agency to change the world.

References

Berliner, D. C., & Biddle, B. J. (1997). *The manufactured crisis: Myths, fraud, and the attack on America's public schools* (collegiate edition). Addison-Wesley.

Bradbury, A. (2020). A critical race theory framework for education policy analysis: The case of bilingual learners and assessment policy in England. *Race, Ethnicity and Education*, *23*(2), 241-260.

Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives, 22(86),*1.

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press.

Friedman, M. (1982). *Capitalism and freedom*. University of Chicago press.

Gibbs, N. P., Pivovarova, M., & Berliner, D. C. (2023). Same tests, same results: Multi-year correlations of ESSA-mandated standardized tests in Texas and Nebraska. *Education Policy Analysis Archives*, *31*.

hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom*. Routledge.

Kezar, A. (2011). Grassroots leadership: Encounters with power dynamics and oppression. *International Journal of Qualitative Studies in Education, 24*(4), 471-500.

Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom*. Beacon Press.

MacLean, N. (2021). How Milton Friedman exploited White supremacy to privatize education. *Institute for New Economic Thinking Working Paper Series*, (161).

Marshall, C., Mitchell, D., & Wirt, F. (1989). *Culture and education policy in the American states.* The Falmer Press.

Schueler, B. E., & Bleiberg, J. F. (2022). Evaluating education governance: Does state takeover of school districts affect student achievement? *Journal of Policy Analysis and Management, 41*(1), 162-192.

Ulrich, W. (1996). A primer to Critical Systems Heuristics for action researchers. *Independent Research Monograph*, Centre for Systems Studies, University of Hull.

Virginia Association of Supervision and Curriculum Development. (2022a, July 27). VASCD Assessment Summit executive summary. Issuu. https://issuu.com/vascd/docs/assessment_summit_executive_summary-8

Virginia Association of Supervision and Curriculum Development. (2022b, October 28). VASCD 2023 General Assembly priorities. Issuu. https://issuu.com/vascd/docs/vascd_2023_general_assembly_priorities