

University of Richmond

UR Scholarship Repository

Honors Theses

Student Research

2019

Re-imagining student achievement : the egalitarian failure of the carrot-stick model in K-12 public education

Michael Johnson
University of Richmond

Follow this and additional works at: <https://scholarship.richmond.edu/honors-theses>



Part of the [Leadership Studies Commons](#)

Recommended Citation

Johnson, Michael, "Re-imagining student achievement : the egalitarian failure of the carrot-stick model in K-12 public education" (2019). *Honors Theses*. 1422.

<https://scholarship.richmond.edu/honors-theses/1422>

This Thesis is brought to you for free and open access by the Student Research at UR Scholarship Repository. It has been accepted for inclusion in Honors Theses by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

**Re-Imagining Student Achievement:
The Egalitarian Failure of the Carrot-Stick Model in K-12
Public Education**

By

Michael Johnson

Honors Thesis

In

Leadership Studies

University of Richmond

Richmond, VA

May 4, 2019

Advisor: Dr. Jessica Flanigan

Abstract

Re-Imagining Student Achievement:

The Egalitarian Failure of the Carrot-Stick Model in K-12 Public Education

Michael Johnson

Committee Members: Dr. Jessica Flanigan, Dr. KimMarie McGoldrick, Dr. Thomas Shields

Achievement gaps between low-income and minority students and their counterparts are among the most pressing education policy issues today. Cash incentivization to students has gained momentum as a potential remedy to reduce disparities in student achievement outcomes. Grading incentive schemes function identically as objections to cash-incentives positioning both within the broader carrot-stick motivation model. Rather than eliminate the widely used grading scheme, however, I conclude that efforts should be redirected towards reducing the saliency of standardized evaluative benchmarks to which incentives are aimed as opposed to reforming the incentives themselves which I refer to as the Revisionary Proposal.

Signature Page for Leadership Studies Honors Thesis

***Re-Imagining Student Achievement:
The Egalitarian Failure of the Carrot-Stick Model in K-12 Public Education***

Thesis presented

by

Michael Johnson

This is to certify that the thesis prepared by *Michael Johnson* has been approved by his/her committee as satisfactory completion of the thesis requirement to earn honors in leadership studies.

Approved as to style and content by:



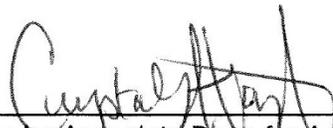
Name, Chair



Name, Member



Name, Member



Senior Associate Dean for Academic Affairs
Jepson School of Leadership Studies

Introduction

Historically, student achievement within U.S. public education has been determined through standardized evaluative metrics such as statewide tests for reading, math and science. Achievement gaps along racial, socioeconomic, and cultural lines are strongly associated with long-term disparities in the future earnings, standards of living, and health indicators between underperforming students and students who succeed.¹ For instance, roughly 15% of low-income 8th graders tested proficient in reading compared to nearly 40% of 8th graders coming from middle-high income backgrounds.² Additionally, the same study finds that the same percentage of low-income students score proficient in math while over 40% of their wealthier counterparts are proficient in math. This trend is supported by a wide body of additional sources.^{3,4} One framework to partially explain the persistence of the income achievement gap is a lack of motivation on behalf of underperforming students. Cash incentivization or giving money to students for meeting or exceeding evaluation standards, has received increased attention as a method to increase student motivation in hopes of reducing the income achievement gap.

The primary research question which I will examine is: to what extent should educators provide cash incentives to reduce achievement gaps? To begin, I will analyze empirical evidence regarding the efficacy of cash incentives in improving short-run student achievement outcomes through various research studies. In the second section, I will build on the work of Michael Sandel and Debra Satz to explore the normative and ethical considerations of cash incentive

¹ “Whither Opportunity?: Rising Inequality, Schools, and Children’s Life Chances.” *Choice Reviews Online* 49, no. 10 (June 1, 2012): 49-5804-49-5804.

² Gorski, Paul. *Reaching and Teaching Students in Poverty: Strategies for Erasing the Opportunity Gap*. New York, NY: Teachers College Press, 2018.

³ Duncan, Greg J., Pamela A. Morris, and Chris Rodrigues. “Does Money Really Matter? Estimating Impacts of Family Income on Young Children’s Achievement with Data from Random-Assignment Experiments.” *Developmental Psychology* 47, no. 5 (September 2011): 1263–79.

⁴ García, Emma, and Elaine Weiss. “Education Inequalities at the School Starting Gate,” n.d., 102.

structures; in particular, their effects on the intrinsic value of education. Next in the third section, I will make the connection between cash and other extrinsic incentives, specifically grades, which both fall under the umbrella term known as the carrot-stick model. This model forms the traditional incentive-reward system in which the carrot represents incentivizing desired values while the stick represents punishing undesirable values.⁵ I argue that educators should reject the current grading system on the same empirical and normative grounds as they should reject cash incentives, not in principle, but within the inherently biased evaluative framework that discourages and de-values those who do not conform to narrowly defined standards. Consequently, I argue that changes to the motivational structure within public education must ultimately follow a re-conceptualization of the evaluative metrics which determine student achievement outcomes by reducing the saliency of standardized testing, also referred to as the Revisionary Proposal. Lastly, I consider and ultimately reject the objection argued by Christopher Knapp which argues that the Revisionary Proposal model would further exacerbate educational disparities. I reject this objection by consulting the work of Jennifer Morton and Christopher Emdin who provide striking evidence regarding the clash between standardized evaluative metrics and the cultural identities of students from marginalized backgrounds.

§1: Cash Incentives and Instrumental Value

The ultimate purposes of education are understood to be both instrumentally and intrinsically valuable. Strict instrumentalists generally point to economic indicators such as employment rates after graduation, income level, and health outcomes to evaluate the success of education for individuals and across localities. The instrumental value of education comprises the first principle of educational egalitarianism. Regarding this principle, two frames of thought

⁵ Hess, Frederick M., and Andrew P. Kelly. *Carrots, Sticks, and the Bully Pulpit: Lessons from a Half-century of Federal Efforts to Improve Americas Schools*. Cambridge, MA: Harvard Education Press, 2011.

exist surrounding the concept of equal opportunity. The first, equality in opportunity, contends that educational resources and opportunities should be distributed equally across students, irrespective of outcome. The second, which I will refer to as equality in outcome, argues that disparities in the provision of educational resources are justified insofar as the short and long-term outcomes are equal for all students. Going forward, I focus on the second conception of educational egalitarianism as the ideal to which domestic public education should achieve. Although the systemic implementation of cash incentives appeared promising, the following empirical research finds that cash incentives are ineffective in improving the achievement outcomes of underperforming students and only slightly effective at best.

Roland Fryer, a Harvard economist, conducted a series of experiments in Dallas, Chicago, and New York in which his research team conducted a study examining the extent to which financial incentives improve student achievement. From 2007 - 2009, he distributed over \$9 million to over 27,000 students across the three cities.⁶ In Dallas, he conducted his experiment on elementary school students, while in New York he tested middle school students and high school students in Chicago. Because the focus of my thesis is concerned with incentives and student achievement at the secondary level, I will only consult the results from his Chicago experiment. Almost 8,000 9th graders participated in his study, half of which were in the treatment group (which received financial incentives) and the other half in the control group. For every A that a student received, he/she would receive \$50 as a financial reward, \$35 for every B, and \$20 for every C. If a student received a D, they would receive no money and a student who

⁶ Fryer, Roland G. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials." *The Quarterly Journal of Economics* 126 (4): 1755–98.

received an F would not receive a financial reward in addition to losing any money they were rewarded within that period.⁷

With this scale, students could earn up to \$250 every 5 weeks or \$2,000 a year. Chicago is extremely compelling in my examination of student achievement as over 88% of students in Chicago public high schools are black or hispanic, 75% are eligible for free and reduced lunch and almost 15% are English language learners. Surprisingly, the results from Fryer's study depicted a very minimal impact of the financial incentive structure in improving the achievement outcomes of the underperforming students in the district. 9th graders witnessed zero effect on Illinois standardized tests which served as the primary outcome metric.⁸ G.P.A. for this group improved very slightly at about .93σ, while student on average earned additional credits translating into approximately half a full course.⁹

The second experimental research design of our study was conducted by Kirabo Jackson and was labeled the AP Incentive Program (APIP).¹⁰ APIP began in 1996 and was initiated in 10 high schools across Dallas, Texas. APIP has produced minimal improvements in academic performance measures, such as AP scores and number of students who enroll in a college or university. According to the study, the number of students scoring above a 1100 on the SAT increased by 30%, and the number of high school students who enroll in a college or university increased by 8%.¹¹ Despite this, Jackson found no statistically significant evidence to suggest that APIP encouraged a greater number of students to take the AP exam or produced higher graduation rates. While the results depict that the biggest demographics resulting in the increase

⁷ Ibid. 1761.

⁸ Ibid. 1757.

⁹ Ibid. 1758.

¹⁰ Jackson, C. Kirabo. "A Little Now for a Lot Later A Look at a Texas Advanced Placement Incentive Program." *Journal of Human Resources* 45, no. 3 (July 1, 2010): 591–639.

¹¹ "Cash for Test Scores." *Education Next*, August 15, 2008. .

of ACT scores and college attendance are primarily from black and Hispanic populations, the study makes minimal reference to low-income populations.

§2: Financial Incentives and Intrinsic Motivation

Although the pursuit of equality in outcome, grounded in instrumental rationales, serves as the first principle of educational egalitarianism, it is not the sole end of egalitarianism. Educational egalitarianism should not just be concerned with the instrumental reasons of economic and employment outcomes but also concerned with the intrinsic purposes of education to develop one's talents, pursue one's interests, and realize one's passions, which are good in themselves; to which no one's rights should be limited. Consequently, the second pillar on which the concept of educational egalitarianism rests is the belief that a child's ability to utilize and develop their talents shouldn't be determined by factors outside of their control such as their socioeconomic status, race, gender, or other arbitrary determinants. Because education is a good which has both instrumental purpose and intrinsic value, even if the previously mentioned studies conclusively determined that financial incentives lead to better academic outcomes for low-income students, it still does not necessarily satisfy the normative question of whether such incentives should be implemented.¹²

As Michael Sandel argues, introducing an instrumentalist mindset through the commodification of a good has moral limits. On this point Sandel presses two main objections: the fairness objection and the corruption objection.¹³ The fairness objection relates to the relationship between the market and inequality. He argues that markets for certain goods

¹² Shields, Liam, Anne Newman, and Debra Satz. "Equality of Educational Opportunity." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2017. Metaphysics Research Lab, Stanford University, 2017.

¹³ Sandel, Michael J. *What Money Can Buy: The Moral Limits of Markets*. New York: Farrar, Straus & Giroux, 2013.

shouldn't be introduced on the basis that they would perpetuate structural and historical inequality. One example he uses to illustrate this objection is the introduction of a market for organs. An introduction of a market for organs, Sandel argues, would have a crowding-out effect of pricing low-income individuals out of the market; therefore leaving them unable to access life-saving transplants. The crowding out effect of this example is aligned with the historical and structural inequalities that limit the poor from accessing equal health care resources leading to disparities in health outcomes across socioeconomic status. Contrary to a market in the sale of organs, a market for academic performance aims to reduce inequality by not only providing money to poor students, but also by incentivizing academic outcomes which could ultimately lead to higher lifetime earnings dependent upon a student's future level of education. Consequently, financial incentives within the market of academic performance for low-income students would be ethically sound according to Sandel's fairness objection.¹⁴

Since the introduction of financial incentives to achieve higher academic outcomes is specifically targeted at low-income populations, the crowding-out effect resulting from the introduction of a market in the transaction of organs is dramatically different from the crowding-out effect of a market for academic performance. Debra Satz offers a further conception of the crowding-out principle as it relates to a market for academic performance. Rather than focusing on pricing individuals out of the market, Satz argues that monetary incentives can crowd out the intrinsic worth of the actions one hopes to encourage.¹⁵ Her contention is consistent with a considerable amount of additional studies which conclude that extrinsic motivations diminish the

¹⁴ Brennan, Jason, and Peter Jaworski. *Markets without Limits: Moral Virtues and Commercial Interests*. New York: Routledge, Taylor & Francis Group, 2016.

¹⁵ Satz, Debra. *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*. New York: Oxford University Press, 2012. 194.

value of any behavior for its own sake.¹⁶ By extrinsic motivation, I am referring to any factor which compels a student to learn, or attempt to learn, beyond his or her passion to learn for the sake of learning. For example, if a student were to read a book about the role of the Federal Reserve for no other reason than a genuine interest in learning about U.S. monetary policy, this would be considered intrinsically motivated. If that same student were to read this book, because of a genuine interest in addition to his desire to receive an A in their macroeconomics course, this would be considered extrinsically motivated.

The monetization of K-12 education could serve to undermine the intrinsic value of learning in addition to the numerous positive externalities associated with school such as genuine friendships, the development of one's character, and other values to a great extent. Satz' conception of the crowding-out effect associated with extrinsic motivations perfectly corresponds with the second objection posed by Sandel; the corruption objection. To this end, financial incentives within the market of academic performance is ethically unjustified according to the crowding-out effect presented by Satz and the corruption objection offered by Sandel. Because of the lack of empirical evidence demonstrating the effectiveness of financial incentives in serving the instrumental purpose of education, in addition to a reduction of the intrinsic value of education as a result of their introduction, I reject the normative question of should cash incentives be used as an alternative means for improving the educational outcomes of low income students.

§3: Implications for Grades

¹⁶ Cerasoli, Christopher P., Jessica M. Nicklin, and Michael T. Ford. "Intrinsic Motivation and Extrinsic Incentives Jointly Predict Performance: A 40-Year Meta-Analysis." *Psychological Bulletin* 140, no. 4 (2014): 980–1008. <https://doi.org/10.1037/a0035661>.

In light of the foregoing argument that the main objections against the provision of cash-based incentives are both normatively and empirically supported, how should educators view other forms of incentives, such as grades and test scores? To review, I established the reduction of the intrinsic value of education for students who learn for cash, and evidence regarding the ineffectiveness of cash-incentives in terms of improving the academic outcomes of low-income students. If the two layered arguments against cash-incentives (intrinsic and instrumental) are accepted, I will argue that grades have similar effects on the education of underperforming students and their learning outcomes. If we are to accept these considerations for cash incentives, we must also accept them for the provision of grades on which our education system has so heavily relied.

Grades, similar to money, serve as extrinsic motivations for students to learn and master subjects required in school. First, consider the crowding-out and corruption concerns raised by Satz and Sandel within the context of grading in K-12 public education. Grades offer students very little immediate tangible benefits as compared to cash. If a low-income student were to receive high marks at the end of the term, this would offer no immediate remedy for many of the challenges accompanying poverty such as food insecurity, home instability, adverse health outcomes and more. In fact, many may argue that the time focused on receiving high grades can take away time and energy which students could dedicate to alleviating some of these issues through working and contributing to the family income. Proponents of this view would argue that grades do not satisfy as extrinsic motivators because grades do not offer immediate benefits to low-income students and therefore do not serve as motivators in themselves.

While the lack of immediate rewards resulting from receiving high grades, and students' knowledge of this, has some degree of validity, it does not provide sufficient reasoning to

conclude that grades do not serve as extrinsic motivators for students because of the long-term extrinsic motivations associated with receiving high grades. Beginning in high school in the K-12 education system, although students may begin earlier, the grades students receive not only are permanently recorded on their academic transcript, but they also are typically aggregated under a common grading scale referred to as Grade Point Average(GPA); although there are some exceptions to this. In immediate terms, if students do not perform at the minimal level determined at the district, state, and federal level, he/she will not graduate and receive their diploma. As previously stated, a high school diploma, or GED, is not sufficient to assure future economic stability for low-income students. Without their high school diploma, their likelihood of being impoverished is even greater.

Beyond the actual completion of their degree, the strength of both their academic transcript and GPA are weighted heavily in college admission decisions and even some places of employment. Since there is the belief and evidence to show that a high school diploma is necessary but insufficient to achieving financial security above the poverty threshold, graduating from a college or university has come to dominate the narrative surrounding the instrumental purpose of education as a mechanism to alleviate structural and generational poverty. As a result, students are motivated to perform well, through grades and extracurriculars, in order for them to have the opportunity to attend an institution of higher learning. If the exponential increase in the number of students attending these institutions were the result of students who simply had the desire to further their education for the sake of learning, this would present a significant hole in my argument. However, self-reported studies have shown that students who wish to attend college or university are significantly motivated by the prospect of better employment outcomes

in the future. For example, in 2016, 55% of students reported that they viewed college as necessary to workforce success.¹⁷

Instrumentally, grades as they have been utilized for decades have not contributed to the elimination of large-scale systematic improvements in the achievement outcomes of students within historically underperforming groups. In fact, grades and the performance metrics by which we evaluate them, have only exacerbated the racial and socioeconomic disparities in the student achievement gap. The National Assessment of Educational Progress (NAEP) currently is the most effective tool for educators and policymakers to measure disparities in student achievement in math and reading proficiency. According to this assessment, racial disparities have steadily decreased since the 1990s and are approximately over 30% smaller than in the 1970s.¹⁸ While this reflects a move in the right direction towards the instrumental end of educational egalitarianism, racial disparities in student achievement continue to persist as Hispanic-white gaps and Black-white gaps range from .5 standard deviations to .9 standard deviations respectively for math. Even more troubling, socioeconomic disparities do not follow the same declining trend as racial disparities throughout the last three decades. Conversely, socioeconomic disparities have increased by over 60% since the 1960s and are now about twice as large as the national racial achievement gap.¹⁹ This trend reflects a retreat from the instrumental aim of educational egalitarianism and is a primary driver behind the growing interest in the implementation of financial incentives to underperforming students who are much more likely to come from lower socioeconomic backgrounds.²⁰

¹⁷ “Why Do Americans Go to College? First and Foremost, They Want Better Jobs.” Washington Post. Accessed September 12, 2018.

¹⁸ “The Educational Opportunity Monitoring Project: Racial and Ethnic Achievement Gaps.” n.d. Accessed March 8, 2019.

¹⁹ Paul, Annie Murphy. n.d. “In Defense of School Testing.” *Time*. Accessed March 8, 2019.

²⁰ “Whither Opportunity? Rising Inequality, Schools, and Children’s Life Chances.” 2012. *Choice Reviews Online* 49 (10): 49-5804-49-5804.

One may reply that while cash serves solely as an incentive, grades perform as incentives and for pedagogical purposes. According to this view, grades serve the purpose of not only motivating students but also as a critical tool to accurately understand a student's level of proficiency for a given subject or topic. However, I do not find this argument sufficient to undermine the similarities among the effect of cash and grading incentives due to neither incentive occurring in isolation of the feedback mechanism required to develop learning.

If educators were to award high achieving students with money or not reward underachieving students with cash without grounding their decision based on some evaluative metric, such as a test, or communicating with students the reward criteria then this objection would hold true and cash incentives would only serve the purpose as an incentive. In accepting this argument however, the same argument would apply to grades. If educators were to award grades without rationalizing a student's grade based off some evaluative metric or without informing students of the grading criteria, to what pedagogical end would grades serve? The contention that cash incentives do not hold pedagogical purposes rests on the misguided assumption that cash incentives exist in isolation of the pedagogical mechanism of assessment and feedback which characterizes the carrot-stick model. Cash incentives, as they have been used for research purposes, would not be instruments which disrupt this traditional mechanism, but merely a new flavor. Money, identical to grades, reflect feedback of a student's performance which comprises the second component of the pedagogical machine. Student B receiving an A on an assignment reflects their proficiency on a given subject to the same extent as if that same student received \$50 assuming that schools used the scale within the Fryer study. Thus, cash and grades should be understood as synonymous not only in terms of their effects on the intrinsic value of education but also in terms of their ability to reflect and act as feedback mechanisms

within the broader carrot-stick pedagogical framework to which the U.S. public education system by-and-large subscribes.

One point of contention, however, could be made regarding the short vs. long term incentivization effects of financial incentives compared to grading incentives. Because of the immediate nature and benefit of financial gains, cash incentives incentivize in the short run more effectively than in the long run. Conversely, grading incentives serve as long-term instrumental motivators to a greater extent than cash incentives due to the deferred benefit received by high achieving students in the form of admission and greater funding for higher education which as we have already established, dramatically improves one's career and economic outlook in the future. To the degree this is valid, it could result in cash and grades having drastically different effects in motivating students toward achievement metrics. However, given the novelty of financial incentives within education, the research has only examined the effects of cash incentives in the short run as evidenced by the Fryer and Jackson studies. Additionally, the empirical data does not support this claim given that the increase in short-run motivation derived from financial incentives were minimal at best.

§4: A Revisionary Proposal: Reducing the Saliency of Evaluative Standards

The aim of the Revisionary Proposal is concerned with conceptualizing how the ideal of educational egalitarianism should be manifested within public education, given our objections to the carrot-stick model, rather than making tangible policy recommendations which may be constrained due to additional factors such as political expediency and economic climate. This revisionary model posits that the egalitarian failure of the carrot-stick incentive model, both intrinsically and instrumentally, is not a result of the incentives themselves but rather their reliance on evaluative standards which necessarily undermine the identities of students at the

margin. Because of this, the revisionary model explicitly states: Educators should re-conceptualize student achievement from the current trajectory which increasingly prioritizes the methods by which students are incentivized to conform to evaluative metrics to a less rigid conception of student achievement that conforms the evaluative metrics to the student. The former, whether through cash or grades, forces underperforming student A either to fundamentally change certain qualities or habits in order to succeed or get left behind both academically and in the future professionally. The latter on the other hand, allows that same student the flexibility to develop their passions, perfect, and market their unique skill set in the absence of strict evaluation parameters without necessarily being punished in terms of future career and economic outcomes.

Measurements of student potential through standardized evaluations are salient within the framework of the public education system and labor market to the extent that they are used to determine a student's eligibility for selective secondary schools, entrance into higher education, employment outcomes, and to evaluate and justify compensation of teachers, schools, and school districts as a whole. Such a heavy reliance forces intense downward pressure on the students to perform academically to the extent that it crowds out many of the interests and potentially lucrative skill sets which are not assessed by standardized tests. If public education officials and administrators were to eliminate grading standards altogether, students would no longer be as intensely encouraged to merely "learn for the test" or extrinsically motivated to learn topics of little interest to them by the prospect of securing a more lucrative career outcome; which would move us closer toward the intrinsic end of educational egalitarianism.

Under the revisionary proposal, suppose students in Fairfax County Public Schools were still assessed using the end of year Standards of Learning (SOL) examination and still received

final grades at the end of each term. While Fairfax County Public Schools would still be permitted to evaluate its students via grading standards strictly to provide feedback and promote learning, it would not allow such standards to be used in the evaluation of their students by institutions of higher education nor employers. This would manifest itself institutionally by prohibiting schools from releasing academic transcripts to universities and restricting employers and universities from requesting academic information such as G.P.A. or standardized exam scores such as the SAT/ACT.

Pedagogically, such a measure would diminish the downward pressure experienced both by teachers to “teach to the test” and students to strictly learn for the test. Within the classroom setting, this would ideally result in greater curricular and pedagogical diversity by schools being able to broaden the expertise and backgrounds of their teachers, allowing them to infuse their own creative spin within core classes of math, language, history, and science in addition to increased elective options. By fostering greater curricular diversity, the genuine and intrinsic interests of a wider segment of students can be reached and nurtured as students who previously lacked interest in the limited and restrictive material of the past would more likely be able to connect their unique interests with potentially greater academic choices available to them. Not only this, but allowing for greater curricular diversity signals to students that their autonomy and identity is at the very least attempting to be valued by the institution which generates trust and greater intrinsic value of education for the student. With greater intrinsic interest in a far greater segment of the student population, innovating new or more intensified incentive schemes would become somewhat unnecessary as students would be increasingly driven simply by their own to learn and hone their skills. Ultimately, by reducing the value that such metrics have on the immediate and long-term outcomes of students, the incentivizing mechanism of student

achievement metrics is diminished as students will no longer fear, or become numb to, the ever-compounding stick associated with underperformance or feel cumberingly pressured toward the elusive carrot of high academic achievement.

Fortunately, many schools and districts within the Northeastern region of the U.S. have already begun implementing models similar to that of which I propose such as the PACE pilot across the state of New Hampshire, Bard Early Colleges in New York, and Montessori Schools. The state of New Hampshire has successfully implemented assessment policies which significantly reduce the saliency of traditional grading practices arguably without threats to student motivation using participation rates as an indicator of motivation. In 2015, New Hampshire implemented the two-year Performance Assessment of Competency Education (PACE) pilot which was the first federally funded assessment program minimizing the salience and reliance on standardized metrics.²¹ This strategy is designed to limit statewide assessments to only twice during primary school (grades 3-8), and once in secondary school (grade 11) while charging instructors throughout the district to collaborate and develop competencies and skills tailored to the local needs and interests of the students; referred to as competency or mastery-based learning. For example, in the Sanborn Regional School District, fourth grade social studies students have traditionally been tested on facts about the NH legislature and governor, however through PACE, students are now students are evaluated on their ability to propose a bill and defend it in a mock hearing.²² The effect of the pilot has been significant; for both reading and math assessments, students' scores improved by more than 3% statewide after the first year.²³ Additionally, reading scores for second grade students increased from 29% to 77% while special

²¹ "Performance Assessment of Competency Education (PACE)." Accessed April 20, 2019.

²² "Goodbye ABCs: How One State Is Moving Beyond Grade Levels and Graded Assessments - EdSurge News." EdSurge, May 16, 2017.

²³ "2015-16 Statewide Assessment Results | NH Department of Education." Accessed April 21, 2019.

education referrals decreased by 21% during this same time period.²⁴ As a result of these improvements, the NH DOE extended the program in 2018 to continue in the direction away from standardized assessments.

§5: Instrumental Threat of the Revisionary Proposal

Up to this point, I have argued that the Revisionary Proposal offers public education officials a method to step closer toward the intrinsic end of educational egalitarianism that the current carrot-stick model fails to reach. However, to what extent does the Revisionary Proposal achieve the instrumental end of educational egalitarianism? Christopher Knapp, a leading educational philosopher, argues that a model similar to the Revisionary Proposal would fail to achieve this instrumental end but potentially exacerbate economic inequities to a greater extent than the carrot-stick incentive structure; which I reject. Contrary to the revisionary proposal which calls for reduced salience of consistent and standardized evaluative metrics, Knapp argues that inconsistent evaluative standards to which incentives are aimed undermines educational egalitarianism rather than promotes it.²⁵ Because students participate in the same academic competition, the argument goes, the standards by which we judge student performance (grades and test scores) should be perfectly consistent across students and institutions if equality and fairness are to be achieved. As a result of relaxing the standards, teachers will be able to give students unfair and objectively unsubstantiated evaluations through similar mechanisms that I proposed as alternatives within the Revisionary Proposal such as letters of recommendation. Additionally, institutions of higher education and employers will be able to make admissions and

²⁴ “Goodbye ABCs: How One State Is Moving Beyond Grade Levels and Graded Assessments - EdSurge News.” EdSurge, May 16, 2017.

²⁵ Knapp, Christopher. 2007. “Assessing Grading.” *Public Affairs Quarterly* 21 (3): 275–294.

hiring decisions similarly based upon unjustified rationales which would not serve to advance the long-term instrumental ends of perpetually underperforming students any more than cash or grading incentive schemes.²⁶

According to this objection presented by Knapp, racial and socioeconomic disparities currently present within education would be exacerbated as the result of the Revisionary Proposal. Increasingly subjective measures will not only grant teachers the ability to use objectively unsubstantiated evaluation criteria, but a lack of standardized evaluation criteria could potentially lead to evaluators relying on considerations of race, ethnicity, gender and other forms of identity categorizations as a means of perpetuating discrimination against those within marginalized groups. This objection, if valid, could potentially suggest that the Revisionary Proposal would threaten the instrumental ends of education to a greater extent than the current carrot-stick model by further placing students within marginalized groups at an even greater disadvantage. I will reject this objection posed by Knapp through consulting evidence of standardization inherently conflicting with the cultural identities of those who typically underperform provided by Jennifer Morton and Christopher Emdin.

In theory, standardized evaluative metrics within public education as argued by Knapp would be morally unobjectionable on egalitarian terms. This is due to the nature of standards to naturally consolidate values and reject contingencies and outliers. Within the context of public education, increasing reliance and saliency of standardized performance measurements has the result of further restricting and narrowing our conception of academic achievement which could account for the growing tension between the increasingly diverse society of the United States and the public education system's continued reliance on narrow evaluative metrics. This is supported

²⁶ Ibid. 280

through Jennifer Morton's work on cognitive dispositions supports my argument rejecting increased evaluative standardization. She suggests that a key factor contributing to income and racial achievement gaps is inherently biased standards which place higher value on cognitive dispositions and devalue non-cognitive dispositions which minority students tend to employ to a greater extent than their counterparts.²⁷ She defines cognitive dispositions as "those that have a basis in or are related to conscious intellectual activity, explicit reasoning, mathematical and verbal ability, and are often measured by standardized tests."²⁸ In contrast, she defines non-cognitive dispositions as those related to, "motivational, social, and emotional dispositions such as self-control, perseverance, emotional stability, self-esteem, etc."²⁹ Based off these definitions, the saliency and prevalence of grading through standardized testing supports the argument that K-12 public education places greater emphasis and offers greater rewards to those who develop their cognitive dispositions, most notably at the secondary level. In fact, students who exhibit high levels of non-cognitive dispositions are frequently punished for behaving outside the norm such as the student who uses creative expressions of music to help herself focus or the student who is genuinely interested in the instruction material yet independently explores a subtopic not specified in the curriculum; both of whom may appear disrespectful to the instructor or as students who do not wish to learn.

Morton supports her claim through research regarding differing parenting styles between middle class and working-class parents conducted by Annette Lareau, a Professor of sociology at the University of Pennsylvania. Utilizing Lareau's research, Morton contends that students from minority and low-income backgrounds further develop non-cognitive dispositions due to

²⁷ Morton, Jennifer M. 2011. "The Non-Cognitive Challenge to a Liberal Egalitarian Education." *School Field* 9 (3): 233–50.

²⁸ Ibid. 239.

²⁹ Ibid.

working class parents assisting their children to develop naturally and with a different relation to authority than children coming from middle class families.³⁰The value differentiation between cognitive and non-cognitive dispositions provided by Morton depicts an inherent bias of the purportedly objective achievement benchmarks which directly contributes to racial and socioeconomic achievement gaps. Within this framework, one reason incentives may fail instrumentally to significantly improve the achievement outcomes of underperforming students is that the ends to which they seek motivate lower income students' conflict with their own identities and upbringing. As a result, a student does not perceive the potential reward resulting from the incentive as sufficient compensation for the potential loss of parts of their cultural identity.

Christopher Emdin, a former high school teacher and the current director of the Institute for Urban and Minority education at Columbia University, provides further evidence to undermine the view presented by Knapp in his book "For White Folks Who Teach in the Hood". He states that, "The noncosmopolitan classroom results in alienating the neo indigenous... This type of classroom forces students to deny their natural abilities and talents and punishes those who refuse to comply while concurrently placing at a disadvantage those who seek to acclimate."

³¹The cultural debasement arising from narrow evaluative metrics tangibly manifests itself through code-switching which both Morton and Emdin refer to throughout their literature. Code-switching, meaning the practice of tailoring one's language, tone of voice, and diction to different environments reflects a student's suppression of their own cultural identity through language in order to conform to the norm set within the educational setting. For many minority,

³⁰ Lareau, Annette. 2011. *Unequal Childhoods: Class, Race, and Family Life*. 2nd ed. University of California Press.

³¹ Emdin, Christopher. *For White Folks Who Teach in the Hood ... and the Rest of Y'all Too Reality Pedagogy and Urban Education*. Beacon Pr, 2017. 110.

lower-income, and ELL students, code-switching becomes necessary in order to achieve academically as students are expected to employ the formal writing and language rules taught throughout school. Not only do formal speech and writing standards unfairly disadvantage those whose language inside the classroom setting sharply contrasts their language outside the classroom but they also convey to the student that their cultural identity, and therefore intelligence, is not as valued as their counterparts within the educational setting.

The appeals to cultural variances by Morton and Emdin provide evidence that the incentive-reward mechanism in K-12 public education not only fails to achieve but is incompatible with the ideal of educational egalitarianism. Educational incentives inherently must be designed toward a particular end that currently exist in the forms of achievement standards. Despite the objection offered by Knapp, the benchmarks education officials set, no matter how consistent, must result in students who excel and students who struggle. Because educational standards reflect and place value on the traditional norms of the majority, which historically has been tailored to white cisgender males, the benchmarks by which we assess student achievement directly conflict with the core identities of immigrants and low-income students of color. Due to the growing multiculturalism within U.S. society, if educational achievement standards and the incentives through which students are expected to achieve them continue to operate within such narrow parameters, the inequalities persistent within achievement gaps will only become exacerbated. if we accept the previous cultural argument that evaluative standards run increasingly counter to the identities and learned skill sets of those who disproportionately underperform, it becomes apparent that their value and salience within access to higher education and the labor market forms the basis for the moral unjustification of student achievement outcomes.

However compelling the evidence presented by Morton and Emdin regarding the cultural clash between marginalized students and increasing evaluative standardization may be, this does not serve as conclusive evidence to invalidate Knapp's objection that educational inequities would be further exacerbated resulting from the Revisionary Proposal. Despite the evidence to suggest the contrary, I will assume that Knapp's argument is valid and that the Revisionary Proposal results in greater inequities within education. Does this serve as reason to unjustify the proposal solely on the grounds that it threatens the instrumental ends of education by worsening economic outcomes for an additional number of students? I argue that the objection presented by Knapp, even if we are to assume it as valid, still does not serve as grounds to unjustify the revisionary proposal due to the greater weight and value gained by students from the intrinsic ends of education. On balance, the extent to which the revisionary proposal would exacerbate educational disparities would need to be great to outweigh the intrinsic ends promoted by the revisionary proposal. Given the significant level seemingly "objective" evaluative metrics already perpetuate educational disparities, it would be unlikely that the Revisionary Proposal would serve to exacerbate them considerably further.

Conclusion

The purpose of this paper is not intended to conclusively advocate for any specific policy recommendations as it would require much greater econometric and statistical analysis. However, the goal of this work is to rigorously grapple with the underlying philosophy and belief formation which ultimately drives policy decisions. In principle, I do not maintain that incentives are universally ineffective, rather I conclude that in practice, incentives are both ineffective and morally objectionable to the extent that they conform students toward ends (evaluative metrics) which diminish a student's intrinsic motivation in addition to inherently disadvantaging certain

students over others. While literature and small-scale alternative models to the current evaluative system exist, such as PACE & Bard, my work highlights the pressing need for greater attention toward systematically re-shaping the beliefs systems undergirding educational techniques rather than the techniques themselves in order to shift what and who is valued by public education. History and ample data have demonstrated the ineffectiveness and moral objections associated with the carrot-stick model manifested by grading incentives and most recently through cash incentives. The growing emphasis to innovate and devise such incentive structures not only highlight the severity of achievement gaps, but more dangerously reflect the ignorance and possible desire of policy makers to undermine, minimize, and threaten the identities and unique experiences of students who do not fit the prototypical mold valued within educational settings. While financial incentives could present a dramatic shift in the way students are educated, education officials must first seriously consider and develop a consistent foundational understanding of the ideal end(s) of education and the role of egalitarianism within it. Until then, administrators and policy makers will continue offering carrots to swaths of students who never desired to eat them.