Vanderbilt Journal of Entertainment & Technology Law

Volume 20
Issue 1 *Issue 1 - Fall 2017*

Article 3

2017

Reshaping Ability Grouping Through Big Data

Yoni H. Carmel

Tammy H. Ben-Shahar

Reshaping Ability Grouping Through Big Data

*Yoni Har Carmel**, *Tammy Harel Ben-Shahar***

ABSTRACT

This Article examines whether incorporating data mining technologies in education can promote equality. Following many other spheres in life, big data technologies that include creating, collecting, and analyzing vast amounts of data about individuals are increasingly being used in schools. This process has already elicited widespread interest among scholars, parents, and the public at large. However, this attention has largely focused on aspects of student privacy and data protection and has overlooked the profound effects data mining may have on educational equality. This Article analyzes the effects of data mining on education equality by focusing on one educational practice—ability grouping—that is already being transformed by educational data mining.

Ability grouping is the practice of separating students into classes or tracks according to their perceived academic abilities. While some educators support the practice, arguing that it helps teachers adjust to the needs of their students, critics argue that ability grouping reinforces educational inequalities. Implicit biases that pervade educational decision-making processes result in the overrepresentation of students from racial and ethnic minorities, and students from poor families, in lower tracks in which they receive inferior education and limited opportunities.

Given the well-documented biases in traditional ability grouping, data-driven ability grouping (DDAG)—the use of algorithms

---

* PhD candidate, Center for Law and Technology, University of Haifa. This research was supported by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation 1716/20. This research was accepted to the 2017 Michigan Young Scholars Conference. The Authors wish to thank Niva Elkin-Koren, Michal Yerushalmy, and the participants at the Education Law Association Conference for their helpful comments.

** Assistant Professor, University of Haifa Faculty of Law and Academic Director, Clinics for Law and Social Change. This work was supported by the German Israel Foundation (GIF) (Research Grant No. I-2484-103.4/2017).

to inform assignment decisions—may be a step in the right direction. However, as this Article demonstrates, the use of data mining technologies for ability grouping creates a host of unique challenges in terms of educational equality.

This Article argues that traditional doctrines of equal protection will be unable to contend with the biases DDAG is likely to create. Instead, this Article offers a novel approach to the legal regulation of DDAG that involves integrating legal and technological expertise and creating equality-sensitive algorithms. The combination between legal and technological solutions can ensure DDAG decreases biases in ability grouping and promotes educational equality.

TABLE OF CONTENTS

I. INTRODUCTION 89
II. THE PRACTICE OF ABILITY GROUPING 94
    A. *What Is Ability Grouping?* 94
    B. *Ability Grouping and Educational Equality* 96
III. DATA-DRIVEN ABILITY GROUPING 103
    A. *Educational Data-Driven Decision-Making* 103
    B. *Can Data-Driven Ability Grouping Reduce Biases?* 107
        1. Discriminatory Attributes 110
        2. Attributes That Correlate with Discriminatory Classifications 110
        3. Representation Within Data 112
        4. Biases in the Design of the Algorithm Itself 113
        5. Why Are Biases Especially Troubling in Data-Driven Ability Grouping? 114
        6. Possible Technological Solutions 115
IV. LEGAL REGULATION OF DATA-DRIVEN ABILITY GROUPING 117
    A. *Challenging Data-Driven Ability Grouping* 117
    B. *Regulating the Design and Implementation of Data-Driven Ability Grouping* 124
V. CONCLUSION 128

I. INTRODUCTION

The practice of grouping students according to their ability affects millions of students in the United States each day.[1] It shapes crucial aspects of their education: the curriculum they study, the resources they receive, the teachers who educate them, and the peers with whom they interact.[2] Critics of ability grouping insist that it reinforces educational inequalities, stratifying students from racial and ethnic minorities and students from poor families to lower tracks in which they receive inferior schooling and limited opportunities.[3] Proponents, on the other hand, argue teaching homogeneous classes is more effective, as it allows teachers to adjust content and pedagogy to the students' needs.[4] All experts concede, however, the importance of ensuring a grouping process that is free from biases and does not aggravate racial or class segregation.[5]

Despite being one of the most controversial issues in education for almost a century, the practice of ability grouping persists and has thrived for the past decade.[6] The resurgence of ability grouping coincides with another momentous change in education—the technological and information revolution.[7] This development, which influences educational practices in myriad ways, already affects ability grouping practices in many schools around the country.[8]

1. TOM LOVELESS, BROWN CTR. ON EDUC. POLICY AT BROOKINGS, THE 2013 BROWN CENTER REPORT ON AMERICAN EDUCATION: HOW WELL ARE AMERICAN STUDENTS LEARNING? 20 (2013), https://www.brookings.edu/wp-content/uploads/2016/06/2013-brown-center-report-web-3.pdf [https://perma.cc/Z2XM-YEES].

2. *Id.*

3. *See, e.g.*, JEANNIE OAKES, KEEPING TRACK: HOW SCHOOLS STRUCTURE INEQUALITY 233, 235, 238 (2d ed. 1985). For a detailed discussion, *see infra* Part II.B.

4. *See* NAT'L EDUC. ASS'N, ACADEMIC TRACKING: REPORT OF THE NEA EXECUTIVE COMMITTEE SUBCOMMITTEE ON ACADEMIC TRACKING 8 (1990); Vivian Yee, *Grouping Students by Ability Regains Favor in Classroom*, N.Y. TIMES (June 9, 2013), http://www.nytimes.com/2013/06/10/education/grouping-students-by-ability-regains-favor-with-educators.html?mcubz=1 [https://perma.cc/E957-KCMN] (describing teachers' positive attitude toward ability grouping as a strategy to cope with student diversity); *see also* Julian R. Betts, *The Economics of Tracking in Education*, *in* HANDBOOK OF THE ECONOMICS OF EDUCATION 341, 341–81 (Eric A. Hanushek et al. eds., 2011) (discussing the challenges in empirical evidence concerning tracking).

5. *See, e.g.*, NAT'L EDUC. ASS'N, *supra* note 4, at 2–4; Betts, *supra* note 4, at 326; Yee, *supra* note 4.

6. *See, e.g.*, LOVELESS, *supra* note 1, at 17 (stating that the frequency of using ability grouping in fourth-grade reading instruction rose from 28 percent in 1998 to 71 percent in 2009).

7. *See, e.g.*, Roger Riddell, *What Trends Are Shaping Ed Tech in 2014?*, EDUC. DIVE (Feb. 6, 2014), http://www.educationdive.com/news/what-trends-are-shaping-ed-tech-in-2014/223048/ [https://perma.cc/Q77J-FV5L].

8. *See* Cristóbal Romero & Sebastián Ventura, *Educational Data Mining: A Review of the State-of-the-Art*, 20 IEEE TRANSACTIONS ON SYS. MAN & CYBERNETICS 1, 9 (2010); Milan

Educational technologies that are increasingly being introduced into schools generate vast amounts of student data, which are collected, mined, and analyzed by algorithms through educational data mining (EDM) techniques.[9] The algorithm outputs can be used for various purposes, including teacher evaluation, improving pedagogy, informing education policy, and the practice that is the focus of this Article: ability grouping.[10]

One of the most interesting questions raised by the use of EDM for ability grouping is whether it will alleviate the biases that plague traditional ability grouping and decrease the overrepresentation of children from minority communities and poor families in the lower educational tracks. These biases have troubled both educators and legal scholars in the past, and while much attention has been devoted to the topic, little progress has been made.[11] The introduction of data-driven ability grouping (DDAG) substantially changes the way grouping is performed and therefore warrants renewed interest in the topic. This Article examines the effects DDAG may have on educational equality, relying on the developing literature pertaining to the ethical and legal ramifications of big data and predictive analytics. Within this body of literature, only sparse attention is given to the educational arena, and the existing research focuses mostly on issues of privacy, data protection, and preventing the monetization of student information.[12] This Article addresses this gap in scholarship and brings together several distinct areas of scholarship—antidiscrimination law, education law, and technology law—the integration of which introduces novel issues of importance for each area of law.

Vukicevic et al., *Grouping Higher Education Students with RapidMiner*, *in* RAPIDMINER: DATA MINING USE CASES AND BUSINESS ANALYTICS APPLICATIONS 185, 185 (Markus Hofmann & Ralf Klinkenberg eds., 2013). For a detailed discussion of this practice see *infra* Part III.A.

9. *See* BARBARA MEANS ET AL., U.S. DEP'T OF EDUC., USE OF EDUCATION DATA AT THE LOCAL LEVEL: FROM ACCOUNTABILITY TO INSTRUCTIONAL IMPROVEMENT 2 (2010), https://www2.ed.gov/rschstat/eval/tech/use-of-education-data/use-of-education-data.pdf [https://perma.cc/U4XN-NRJJ]; Vukicevic et al., *supra* note 8, at 185.

10. *See* Romero & Ventura, *supra* note 8, at 3–11.

11. *See* Mary Cipriano-Walter, *Falling off the Track: How Ability Tracking Leads to Intra-School Segregation*, 41 T. MARSHALL L. REV. 25, 47 (2016); Anthony D. Greene, *Tracking Work: Race-Ethnic Variation in Vocational Course Placement and Consequences for Academic and Career Outcomes*, 1 INT'L J. EDUC. STUD. 9, 11 (2014).

12. *See generally* Jules Polonetsky & Omer Tene, *The Ethics of Student Privacy: Building Trust for Ed Tech*, 21 INT'L REV. INFO. ETHICS 25 (2014); Elana Zeide, *Student Privacy Principles for the Age of Big Data: Moving Beyond FERPA and FIPPS*, 8 DREXEL L. REV. 339 (2016); Joel Reidenberg et al., *Privacy and Cloud Computing in Public Schools*, CTR. L. & INFO. POL'Y, Dec. 2013, http://ir.lawnet.fordham.edu/clip/2.

This Article argues DDAG offers significant promise by potentially removing prejudice from educational decisions, thus offsetting implicit biases that teachers may unwittingly hold. A recent study examined an algorithm-based system called the Education Value-Added Assessment System (EVAAS), used for assigning students to different tracks in eighth-grade mathematics.[13] The study found that the algorithm assigned students to a higher track when the students otherwise would not have been identified as suitable for the track, thus increasing the proportion of children from racial minorities and low socioeconomic status in the higher track.[14]

Despite the promise it extends, DDAG creates a host of unique challenges in terms of equality of opportunity. Studies on data mining and predictive analytics in other domains such as crime prevention, banking, and insurance suggest that instead of eliminating social biases, algorithms recreate them.[15] To generate predictions, algorithms use historical datasets from which they infer the attributes of potential criminals, potential reckless drivers, or debtors who are likely to fail to pay their debt.[16] When historical datasets are racially biased, the algorithm's decisions simply mirror those biases.[17]

Additionally, algorithms rely on what data they have.[18] Students from a privileged background have better access to digital devices outside of school, meaning they will likely register more entries into the system and record more academic interaction and task engagement.[19] These additional entries consequently have a positive

---

13. *See* Shaun M. Dougherty et al., *Middle School Math Acceleration and Equitable Access to Eighth-Grade Algebra: Evidence from the Wake County Public School System*, 37 EDUC. EVALUATION & POL'Y ANALYSIS 80S, 81S (2015), http://journals.sagepub.com/doi/pdf/10.3102/0162373715576076 [https://perma.cc/P2MG-DFM8]. According to a report on the company's website, EVAAS is widely used to place students in eighth-grade algebra. *Expanding Eighth-Grade Algebra Participation*, SAS INST. INC., https://www.sas.com/en_us/customers/wake-forest-rolesville.html [https://perma.cc/22PW-AM8C] (last visited Sept. 3, 2017); *see also* SAS, EVAAS® FOR K-12: STATISTICAL MODELS 3 (2016), https://evaas.sas.com/support/EVAASStatisticalModels.pdf [https://perma.cc/DV6Z-RQZT]. For further discussion of EVAAS, see *infra* Part III.B.

14. *See* Dougherty et al., *supra* note 13, at 81S. The study also found that the rates of success did not decline subsequently. *See id.* at 93S.

15. *See* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 674 (2016). For a detailed discussion, see *infra* Part III.B.

16. *See* Barocas & Selbst, *supra* note 15, at 680.

17. *See* FAISAL KAMIRAN & INDRE ŽLIOBAITE, *Explainable and Non-Explainable Discrimination in Classification*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES 155, 156 (Bart Custers et al. eds., 2013).

18. *See* Barocas & Selbst, *supra* note 15, at 674.

19. *Cf. id.* at 686 n.57.

effect on the algorithm system's outputs about those students.[20] Students from a privileged background are also considerably more digitally literate, which results in better functioning in a digital environment.[21] These disparities do not reflect an actual gap in academic ability; therefore, they cause the algorithm's prediction to be biased against children from poor families or racial minorities.[22]

Finally, data-driven decision-making (DDDM) may create new classes of children who are disadvantaged. Although law is primarily concerned with biases against students belonging to groups that are historically socially excluded, such as racial minorities or immigrants, the Authors contend that algorithmic decision-making may create new groups that are systematically unfairly disadvantaged. If, for some reason, children who are color blind or who engage in after-school sports are less likely to succeed on computerized tasks and, therefore, the algorithmic predictions are less favorable for them, DDDM may be detrimental to their educational prospects, and they may be discriminated against in ability grouping processes.

In at least one sense, the fact that algorithmic decision-making is widely believed to be scientific and objective makes biases in it worse than biases in traditional decision-making. Inequalities that result from DDDM may be perceived as inevitable or justified. This problem is especially challenging in the educational domain, wherein assignment decisions reflect—and influence—children's abilities.[23] By determining the curriculum a child is taught, the skills she develops, the peers she interacts with, the expectations teachers have of her, and the expectations she has of herself, the algorithm's prediction is self-fulfilling.

In light of these concerns it seems reasonable to turn to law to ensure DDAG decreases biases and overrepresentation of minorities in the lower tracks. This, the Authors argue, cannot be achieved through the traditional doctrines concerning equal protection. The existing equal protection doctrines have been largely ineffective in challenging traditional ability grouping practices and, we argue, are even less likely to appropriately address the challenges of DDAG.[24]

The solution instead lies in the combination of technological solutions and legal regulation, both of which should be performed at the stage of the design and use of algorithms. In traditional methods

---

20. *See* Jonas Lerman, Response, *Big Data and Its Exclusions*, 66 STAN. L. REV. ONLINE 55, 56 (2013).

21. *Id.* at 57.

22. *Cf.* Barocas & Selbst, *supra* note 15, at 673.

23. *See* Dougherty et al., *supra* note 13, at 81S.

24. *See infra* Part IV.

of ability grouping performed by humans, it is almost impossible to impose rules concerning which data to use (and which to disregard). It is also extremely difficult to consciously assign a specific weight to each piece of information.[25] Teachers use student grades, tests, and their own impressions to make decisions.[26] Biases are (one hopes) subconscious and unintended, but are hard to avoid. By using algorithms, on the other hand, decision-making is structured and technologically determined. Designers can define which attributes are taken into consideration, which are disregarded, and the weight the algorithm should assign to each. Algorithmic decision-making even enables programmers to determine the desired end result in terms of group representation.[27] Therefore, involvement of legal and normative considerations at the design stage can be effective in decreasing biases and improving outcomes in terms of equality.[28]

Information scientists have already begun seeking technological tools to reduce biased decision-making.[29] These include removing suspect attributes (such as race or gender)[30] and attributes that correlate with suspect attributes (zip code may correlate with race, for example)[31] from the datasets. Another possibility involves manipulating historical datasets from which algorithms learn their predictions by recognizing and correcting biased decisions.[32] Additionally, algorithms may be able to reshape grouping entirely by, for example, replacing the traditional criterion of academic performance with other attributes previously impossible to ascertain, such as different learning styles. This kind of grouping may promote

---

25. Research from a completely different context shows that judicial instruction to jurors to ignore inadmissible evidence does not eliminate the impact of that evidence on jury verdicts. *See* Nancy Steblay et al., *The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis*, 30 LAW & HUM. BEHAV. 469, 469 (2006).

26. John N. Drowatzky, *Tracking and Ability Grouping in Education*, 10 J.L. & EDUC. 43, 45–47 (1981). These decisions are also affected by parental involvement. *See* Elizabeth L. Useem, *Middle Schools and Math Groups: Parents' Involvement in Children's Placement*, 65 SOC. EDUC. 263, 275 (1992).

27. Sicco Verwer & Toon Calders, *Introducing Positive Discrimination in Predictive Models*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES, *supra* note 17, at 263, 263.

28. *Id.*

29. *See* Barocas & Selbst, *supra* note 15, at 716.

30. Toon Calders & Indre Žliobaite, *Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES, *supra* note 17, at 43, 45.

31. *Id.* at 47; Barocas & Selbst, *supra* note 15, at 691–92.

32. Sara Hajian & Josep Domingo-Ferrer, *Direct and Indirect Discrimination Prevention Methods*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES, *supra* note 17, at 241, 247; Verwer & Calders, *supra* note 27, at 255, 263.

the goal of facilitating effective teaching without creating racial and class segregation.

Technological solutions such as these, however, involve numerous normative decisions that cannot be divorced from legal doctrine. It requires, for example, determining which classes are protected, whether unequal outcomes constitute an actionable wrongdoing, and whether affirmative action is permissible. These legal issues, among others, must inform the algorithm designers' decisions. Together, technological and legal regulation can potentially improve the ability grouping process and promote educational equality.

This Article unfolds as follows: Part II describes the current practice of ability grouping and the biases that pervade it. Part III introduces DDAG, explains how it is performed, and discusses whether it is likely to decrease biases in ability grouping. Part IV discusses the existing evidence on biases in predictive analytics and also assesses some possible technological solutions. Part V addresses the role the law can play to ensure that DDAG is used to promote equal educational opportunity and then briefly concludes.

II. THE PRACTICE OF ABILITY GROUPING

A. What Is Ability Grouping?

One of the greatest challenges of comprehensive education lies in the wide variation of students' innate abilities, knowledge, and learning styles.[33] Providing instruction suitable for all students—sufficiently challenging for them but not overwhelming—is an excruciating task. Faced with this challenge, many education systems divide students into groups based on their academic ability, thus decreasing heterogeneity in the classroom.[34] Teachers are then

---

33. *See* OAKES, *supra* note 3, at 3.

34. *See id.* ("Tracking is the process whereby students are divided into categories so that they can be assigned in groups to various kinds of classes."); Patrick Akos et al., *Early Adolescents' Aspirations and Academic Tracking: An Exploratory Investigation*, 11 PROF. SCH. COUNSELING 57, 58 (2007) (describing a tracking policy as involving a school organization structure that increases the homogeneity of instructional groups by stratifying students by curriculum standards, educational and career goals, or ability); Adam Gamoran et al., *An Organizational Analysis of the Effects of Ability Grouping*, 32 AM. EDUC. RES. J. 687, 688 (1995) ("Ability grouping is the practice of dividing students for instruction according to their purported capacities for learning.").

In its widest interpretation, ability grouping includes programs for the gifted on the one hand and placement in special education on the other. *See* Akos et al., *supra*, at 58. While the Authors do not refer to these further in this Article, research has found biases in these decisions too; therefore, some of the discussion applies to these cases. *See* Jesse O. Erwin & Frank C.

able to match the content, pace, and complexity of their classes to their students, who are all, supposedly, more or less at the same ability level.[35]

Ability grouping can take various forms that differ on several dimensions: it can be flexibly performed ad hoc within a diverse classroom for a specific task and dissolve immediately after completion of the task.[36] Conversely, ability grouping can be fixed when students are assigned to separate classes, tracks, or schools from which there is little possibility to move.[37] A second and related dimension concerns the scope of separation. In some cases, grouping entails assignment to completely different schools or tracks in which no mixed ability learning or social interaction takes place.[38] In other cases, schools are comprehensive and ability grouping is used only for specific courses.[39]

Another difference among types of ability grouping policy concerns the age at which ability grouping takes place. In Germany and Austria, for example, students are tracked into separate schools

---

Worrell, *Assessment Practices and the Underrepresentation of Minority Students in Gifted and Talented Education*, 30 J. PSYCHOEDUCATIONAL ASSESSMENT 74, 74–75 (2012) (demonstrating the underrepresentation of minority children in gifted programs); Donna Y. Ford, *The Underrepresentation of Minority Students in Gifted Education: Problems and Promises in Recruitment and Retention*, 32 J. SPECIAL EDUC. 4, 4 (1998) (demonstrating also the underrepresentation of minority children in gifted programs); Robert A. Garda, Jr., *The New IDEA: Shifting Education Paradigms to Achieve Racial Equality in Special Education*, 56 ALA. L. REV. 1071, 1090 (2005) (discussing racial biases in placement of children in special education).

35. JUDITH IRESON & SUSAN HALLAM, ABILITY GROUPING IN EDUCATION 152 (2001); NAT'L EDUC. ASS'N, *supra* note 4, at 8; Garry Hornby et al., *Policies and Practices of Ability Grouping in New Zealand Intermediate Schools*, 26 SUPPORT FOR LEARNING 92, 92 (2011).

36. *See* Saiying Steenbergen-Hu et al., *What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K–12 Students' Academic Achievement: Findings of Two Second-Order Meta-Analyses*, 86 REV. EDUC. RES. 849, 850 (2016).

37. *Id.*; *see also* Maureen T. Hallinan et al., *Ability Grouping and Student Learning*, 6 BROOKINGS PAPERS EDUC. POL'Y 95, 103 (2003). Assignment to lower-track courses can also cause a "locking out" effect when assignment to higher-level courses is conditioned on prerequisite course completion. *See* NAT'L EDUC. ASS'N, *supra* note 4, at 9; George Ansalone, *Schooling, Tracking, and Inequality*, 7 J. CHILD. & POVERTY 33, 42 (2001). Some researchers use the term "tracking" to denote ability grouping that involves completely separate and relatively fixed classification. *See* OAKES, *supra* note 3, at 3; Akos et al., *supra* note 34, at 58; Gamoran et al., *supra* note 34, at 690. Other researchers use the two terms "tracking" and "ability grouping" interchangeably. *See* Steenbergen-Hu et al., *supra* note 36, at 850. This Article uses the more general term "ability grouping."

38. Volker Meier & Gabriela Schütz, *The Economics of Tracking and Non-Tracking* 4 (Ifo Inst. for Econ. Research at the Univ. of Munich, Working Paper No. 50, 2007).

39. Robert E. Slavin, *Ability Grouping and Student Achievement in Elementary Schools: A Best-Evidence Synthesis*, 57 REV. EDUC. RES. 293, 295 (1987).

at the early age of fourth grade, whereas other educational systems are comprehensive until the higher grades.[40]

There is no necessary link between ability grouping and curriculum differentiation, so ability grouping may vary based on the various content and skills students are exposed to in their group.[41] For example, when grouping first-grade children according to their reading ability for tutoring sessions, the goal is to promote their reading skills. Although there may be some differences in the reading material children are given, the curriculum is ultimately the same and the pedagogical aims are identical. The only major difference lies in the pace of progress. Other instances of ability grouping involve completely different curricula and educational goals wherein students acquire different skills and capacities.[42]

Ability grouping in the United States, like other issues in education policy, varies according to local policy.[43] As a rule, however, most US students attend comprehensive schools. Ability grouping does not, therefore, usually involve extreme separation and happens either within classrooms (in elementary schools for reading and math) or by course assignment in middle schools and high schools.[44]

B. Ability Grouping and Educational Equality

For over three decades, education researchers have fiercely debated the effectiveness of ability grouping, and the jury is still out on its effects for educational attainment.[45] While some studies have

40. *See* Meier & Schütz, *supra* note 38, at 2. Ability grouping can sometimes transcend the classic division into grades with cross-grade grouping—an option to address high-ability students' need for accelerated teaching in certain topics. James A. Kulik & Chen-Lin C. Kulik, *Meta-Analytic Findings on Grouping Programs*, 36 GIFTED CHILD Q. 73, 75 (1992).

41. Janet Ward Schofield, *International Evidence on Ability Grouping with Curriculum Differentiation and the Achievement Gap in Secondary Schools*, 112 TCHRS. C. REC. 1492, 1496 (2010).

42. *Id.*

43. SAMUEL ROUNDFIELD LUCAS, TRACKING INEQUALITY: STRATIFICATION AND MOBILITY IN AMERICAN HIGH SCHOOLS 158 (1999).

44. *See id.* at 20; Sean Kelly, *The Contours of Tracking in North Carolina*, 90 HIGH SCH. J. 15, 25 (2007); Kulik & Kulik, *supra* note 40, at 75.

45. *Compare* Esther Duflo, Pascaline Dupas & Michael Kremer, *Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya*, 101 AM. ECON. REV. 1739, 1740 (2011) (finding large and lasting positive effects on the achievement of high- and low-achieving students alike), *with* Robert E. Slavin, *Ability Grouping in the Middle Grades: Achievement Effects and Alternatives*, 93 ELEMENTARY SCH. J. 535, 535 (1993) (reviewing twenty-seven studies concerning middle school and finding almost no difference between students who were grouped according to ability and those who studied in heterogeneous classes).

found positive effects for students studying in homogeneous classes,[46] others have found few or no such effects.[47] Various studies suggest grouping benefits students on higher tracks, whereas students on the lower tracks have no comparable gains[48] and are even disadvantaged by the separation.[49]

Overshadowing the debate on ability grouping effectiveness is the concern that it creates and worsens educational inequality.[50] Two related questions arise here: first, whether ability grouping contributes to widening the gap between high-ability and low-ability

---

46. *See* KELLY PUZIO & GLENN COLBY, SOC'Y FOR RESEARCH ON EDUC. EFFECTIVENESS, THE EFFECTS OF WITHIN CLASS GROUPING ON READING ACHIEVEMENT: A META-ANALYTIC SYNTHESIS (2010) (finding a positive effect for within-class grouping in reading instruction); Duflo, Dupas & Kremer, *supra* note 45, at 1740; Yiping Lou et al., *Within-Class Grouping: A Meta-Analysis*, 66 REV. EDUC. RES. 423, 451 (1996) (finding that within-class ability grouping improved academic achievement); Lynn M. Mulkey et al., *The Long-Term Effects of Ability Grouping in Mathematics: A National Investigation*, 8 SOC. PSYCHOL. EDUC. 137, 137 (2005) (noting that ability grouping in mathematics has persistent instructional benefits for all students); Courtney A. Collins & Li Gan, *Does Sorting Students Improve Scores? An Analysis of Class Composition* 1 (Nat'l Bureau of Econ. Research, Working Paper No. 18848, 2013), http://www.nber.org/papers/w18848 [https://perma.cc/QP4B-EEGL] (noting the performance of both high- and low-performing students significantly improved in math and reading).

47. *See* Slavin, *supra* note 45, at 535; Robert E. Slavin, *Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis*, 60 REV. EDUC. RES. 471, 471 (1990) (reviewing twenty-nine studies examining the effect of ability grouping on achievement in secondary schools, and finding zero effect); *see also* Julian R. Betts & Jamie L. Shkolnik, *The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools*, 19 ECON. EDUC. REV. 1, 1 (2000) (finding no overall effect of formal grouping policies on student achievement).

48. Adam Gamoran et al., *Upgrading High School Mathematics Instruction: Improving Learning Opportunities for Low-Achieving, Low-Income Youth*, 19 EDUC. EVALUATION & POL'Y ANALYSIS 325, 325 (1997) (stating that growth in student achievement in college-preparatory classes is significantly larger than in general track classes); Chen-Lin C. Kulik & James A. Kulik, *Effects of Ability Grouping on Secondary School Students: A Meta-Analysis of Evaluation Findings*, 19 AM. EDUC. RES. J. 415, 425 (1982); Carolyn M. Shields, *A Comparison Study of Student Attitudes and Perceptions in Homogeneous and Heterogeneous Classrooms*, 24 ROEPER REV. 115, 119 (2002) (finding that grouping benefits students with high ability in terms of both academic achievement and attitudes concerning themselves and school).

49. *See* NAT'L RES. COUNCIL, HIGH STAKES: TESTING FOR TRACKING, PROMOTION, AND GRADUATION 102 (Jay P. Heubert & Robert M. Hauser eds., 1999); Estela Godinez Ballón, *Racial Differences in High School Math Track Assignment*, 7 J. LATINOS & EDUC. 272, 272 (2008); Robert L. Linn, *Assessments and Accountability*, 29 EDUC. RESEARCHER 4, 14 (2000); Christy Lleras & Claudia Rangel, *Ability Grouping Practices in Elementary School and African American/Hispanic Achievement*, 115 AM. J. EDUC. 279, 279 (2009) (stating that the progress of students in low-achieving reading groups decreases through the years, thus enlarging the achievement gap); Frances R. Spielhagen, *Algebra for Everyone? Student Perceptions of Tracking in Mathematics*, 5 MIDDLE GRADES RES. J. 213, 214 (2010).

50. Eric A. Hanushek & Ludger Wößmann, *Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries* 10 (Nat'l Bureau of Econ. Research, Working Paper No. 11124, 2005).

students, and second, how ability grouping influences students from disadvantaged families and minority groups.

Most writers on grouping have concluded grouping students by academic performance typically contributes to widening the achievement gap between high-level and low-level classes over time, even after controlling for initial differences in ability.[51] Ability grouping leads to inequality in educational resources: students on lower tracks, despite their need for extra help, tend to receive *fewer* resources than students on the higher tracks,[52] are taught by less experienced teachers,[53] and suffer from negative peer effects.[54] Further, research suggests students on lower tracks are exposed to curricula and learning experiences inferior to those offered on higher tracks.[55] Instruction in low-ability classes tends to be comprised of low-level pedagogy—focusing on isolated bits of information and

---

51. *See* Michael Becker et al., *Is Early Ability Grouping Good for High-Achieving Students' Psychosocial Development? Effects of the Transition into Academically Selective Schools,* 106 J. EDUC. PSYCHOL. 555, 556 (2014); Adam Gamoran & Mark Berendes, *The Effects of Stratification in Secondary Schools: Synthesis of Survey and Ethnographic Research,* 57 REV. EDUC. RES. 415, 415 (1987); Adam Gamoran & Robert D. Mare, *Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?,* 94 AM. J. SOC. 1146, 1146 (1989); Hallinan et al., *supra* note 37, at 104; Thomas B. Hoffer, *Middle School Ability Grouping and Student Achievement in Science and Mathematics,* 14 EDUC. EVALUATION & POL'Y ANALYSIS 205, 223 (1992); Joseph Murphy & Philip Hallinger, *Equity as Access to Learning: Curricular and Instructional Treatment Differences,* 21 J. CURRICULUM STUD. 129, 129 (1989); James E. Rosenbaum, *Social Implications of Educational Grouping,* 8 REV. RES. EDUC. 361, 368 (1980); Alan C. Kerckhoff, *Effects of Ability Grouping in Secondary School in Great Britain* 30 (Nat'l Child Dev. Study, Working Paper No. 9, 1986).

52. *See* Karl L. Alexander et al., *Curriculum Tracking and Educational Stratification: Some Further Evidence,* 43 AM. SOC. REV. 47, 64 (1978).

53. *See* JOAN E. TALBERT & MICHELE ENNIS, STANFORD CTR. FOR RESEARCH ON THE CONTEXT OF TEACHING, TEACHER TRACKING: EXACERBATING INEQUALITIES IN THE HIGH SCHOOL 16 (1990); Merrilee K. Finley, *Teachers and Tracking in a Comprehensive High School,* 57 SOC. EDUC. 233, 242 (1984); Richard Harker & Peter Tymms, *The Effects of Student Composition on School Outcomes,* 15 SCH. EFFECTIVENESS & SCH. IMPROVEMENT 177, 179–80 (2004).

54. *See* Yehezkel Dar & Nura Resh, *Classroom Intellectual Composition and Academic Achievement,* 23 AM. EDUC. RES. J. 357, 357 (1986). Some studies show that grouping students by ability results in a reduction of peer effects in general. Ron Zimmer, *A New Twist in the Educational Tracking Debate,* 22 ECON. EDUC. REV. 307, 307 (2003); Ron W. Zimmer & Eugenia F. Toma, *Peer Effects in Private and Public Schools Across Countries,* 19 J. POL'Y ANALYSIS & MGMT. 75, 75 (2000). Others, however, show that grouping creates a resource-rich environment for high-level students and deprives students on the lower tracks of an important classroom resource—namely, the positive input of high-ability peers. *See* Dar & Resh, *supra,* at 357; Adam Gamoran & Martin Nystrand, *Tracking, Instruction and Achievement,* 21 INT'L J. EDUC. RES. 217, 217 (1994); Sean Kelly & William Carbonaro, *Curriculum Tracking and Teacher Expectations: Evidence from Discrepant Course Taking Models,* 15 SOC. PSYCHOL. EDUC. 271, 273 (2012); Mieke Van Houtte, *Tracking Effects on School Achievement: A Quantitative Explanation in Terms of the Academic Culture of School Staff,* 110 AM. J. EDUC. 354, 359 (2004).

55. *See* Gamoran et al., *supra* note 34, at 692.

workbook usage[56]—that does not develop the students' critical and abstract thinking skills.[57] Being placed on low academic tracks is also related to higher dropout rates,[58] and student misbehavior was disciplined more severely when it occurred on the lower tracks.[59]

Another long-term negative effect associated with being placed on a lower academic track concerns labeling. Grouping dictates teachers' expectations from students and also students' self-expectations.[60] These expectations not only affect students' self-esteem but also influence their actual academic performance.[61] In most cases, once students are placed on a lower academic track in the early grades, they remain there through high school, where the differences between tracks become more pronounced.[62] Students assigned to lower-track courses often find themselves "locked out" of higher-level courses that set conditions for enrollment.[63] As a result, gaps in student achievement tend to widen as students progress through middle and high school, reflecting both the differentiated curriculum and the vast differences in learning opportunities associated with participation in the honors and college preparatory programs available in those schools.[64] This evidence raises grave concerns that instead of improving the academic abilities and attainment of students with lower abilities and investing extra resources in them, ability grouping in fact further disadvantages those students.

The findings are all the more troubling since considerable research shows ability grouping is also detrimental to the educational opportunities of children from poor backgrounds and racial

---

56. *See id.*

57. *See* OAKES, *supra* note 3, at 76.

58. *See* Daniel J. Losen, *Silent Segregation in Our Nation's Schools*, 34 HARV. C.R.-C.L. L. REV. 517, 522 (1999); Jacob Werblow et al., *On the Wrong Track: How Tracking Is Associated with Dropping out of High School*, 46 EQUITY & EXCELLENCE EDUC. 270, 272 (2013).

59. *See* MARY HAYWOOD METZ, CLASSROOMS AND CORRIDORS: THE CRISIS OF AUTHORITY IN DESEGREGATED SECONDARY SCHOOLS 106 (1978).

60. Alexander et al., *supra* note 52, at 60; Harker & Tymms, *supra* note 53, at 179.

61. Aaron M. Pallas et al., *Ability-Group Effects: Instructional, Social, or Institutional?*, 67 SOC. EDUC. 27, 28 (1994) (noting that students in high-ability classes typically are exposed to a more positive learning environment, in terms of attitude, aspirations, and self-esteem, than those in low-ability classes); *see also* Losen, *supra* note 58, at 522.

62. *See* Alexander et al., *supra* note 52, at 56; Doug Archbald & Elizabeth N. Farley-Ripple, *Predictors of Placement in Lower Level Versus Higher Level High School Mathematics*, 96 HIGH SCH. J. 33, 48 (2012); Sean Kelly, *The Black-White Gap in Mathematics Course Taking*, 82 SOC. EDUC. 47, 61 (2009).

63. *See* NAT'L EDUC. ASS'N, *supra* note 4, at 9; Ansalone, *supra* note 37, at 42.

64. *See* Roslyn Arlin Mickelson & Anthony D. Greene, *Connecting Pieces of the Puzzle: Gender Differences in Black Middle School Students' Achievement*, 75 J. NEGRO EDUC. 34, 34 (2006).

minorities.[65] These students are heavily overrepresented in lower tracks, whereas students from privileged backgrounds tend to be assigned in higher proportions to higher tracks.[66]

The fact that children from disadvantaged backgrounds are overrepresented in lower tracks can be attributed to one of two causes. The first, pregrouping causes, are the social circumstances that render children from marginalized groups less equipped for school. Individuals from disadvantaged groups tend to have less nurturing environments, which results in diminished abilities when they enter school.[67] The grouping process at school merely reflects the social inequality. The second possible cause for overrepresentation lies within the process of ability grouping itself—racial and class biases held by educators result in students who could have been successful on the higher tracks being assigned to lower tracks.[68]

Clearly, these two causes are not mutually exclusive. Longstanding social inequality is certainly to blame for inequalities in educational capabilities for children of different social groups. However, there is also evidence that educational decision-making is deeply afflicted with racial and class biases. This Article focuses on the second cause—namely, biases in decision-making—and examines whether the use of EDM coupled with appropriate legal regulation is likely to overcome biases.

Well-documented evidence points to bias in traditional educational decision-making against racial minorities,[69] children of low social class,[70] and female students.[71] Though teachers may be wholly unaware of their biases, they tend to judge equally qualified students from racial minorities as less academically and socially

65. *See, e.g.*, Ansalone, *supra* note 37, at 33; Losen, *supra* note 58, at 519; Hanushek & Wößmann, *supra* note 50, at 13.

66. *See* Ansalone, *supra* note 37, at 39–40; Cipriano-Walter, *supra* note 11, at 27; Greene, *supra* note 11; Losen, *supra* note 58, at 517–18; Jeannie Oakes, *Two Cities' Tracking and Within-School Segregation*, 96 TCHRS. C. REC. 681 (1995).

67. JOHN ERMISCH, MARKUS JÄNTTI & TIMOTHY M. SMEEDING, FROM PARENTS TO CHILDREN: THE INTERGENERATIONAL TRANSMISSION OF ADVANTAGE 181 (2012); ANNETTE LAREAU, UNEQUAL CHILDHOODS: CLASS, RACE, AND FAMILY LIFE 13 (ed. 2003).

68. *See* OAKES, *supra* note 3, at 247; JEANNIE OAKES & AMY STUART WELLS, BEYOND THE TECHNICALITIES OF SCHOOL REFORM: POLICY LESSONS FROM DETRACKING SCHOOLS 23 (1996).

69. *See* Hallinan et al., *supra* note 37, at 103; Terry Kershaw, *The Effects of Educational Tracking on the Social Mobility of African Americans*, 23 J. BLACK STUD. 152, 160 (1992).

70. George Ansalone, *Keeping on Track: A Reassessment of Tracking in the Schools*, 7 RACE GENDER & CLASS EDUC. 108, 112 (2000).

71. Kar L. Alexander & Edward L. McDill, *Selection and Allocation Within Schools: Some Causes and Consequences of Curriculum Placement*, 41 AM. SOC. REV. 963, 973 (1976) (finding that gender influences ability grouping decisions after controlling for ability); Caroline Hodges Persell, Sophia Catsambis & Peter W. Cookson, Jr., *Differential Asset Conversion: Class and Gendered Pathways to Selective Colleges*, 65 SOC. EDUC. 208, 221 (1992).

competent than nonminority students, thus underestimating the students' actual academic abilities.[72] These biases pervade all spheres of schooling. African American children, for example, are more likely to be disciplined for misconduct that white children could get away with—and to suffer more severe punishments for similar behavior.[73] Biases are also connected to decisions concerning assignment to special education[74]: African American children are three times more likely to be found in need of special education when diagnosis of the disability involves subjective teacher evaluations.[75] Such biases do not come forth for more "objective" disabilities such as sensory or physical.[76] Further, while the legal treatment of discrimination and attitudes in society regarding racial equality have developed significantly since these topics were first studied, implicit biases still pervade decision-making.[77]

Students from socially disadvantaged backgrounds are also overrepresented in low-ability tracks owing to differences between affluent and disadvantaged families in parental involvement.[78] Poor parents or parents belonging to minority groups are less likely to challenge assignment decisions than middle- and upper-class parents.[79] Affluent parents are more involved in educational decisions and are more assertive; therefore, affluent parents are more effective

72. *See* Regina Cecelia McCombs & Judith Gay, *Effects of Race, Class, and IQ Information on Judgments of Parochial Grade School Teachers*, 128 J. SOC. PSYCHOL. 647, 647 (1988); La Vonne I. Neal et al., *The Effects of African American Movement Styles on Teachers' Perceptions and Reactions*, 37 J. SPECIAL EDUC. 49, 55 (2003); Felicia R. Parks & Janice H. Kennedy, *The Impact of Race, Physical Attractiveness, and Gender on Education Majors' and Teachers' Perceptions of Student Competence*, 37 J. BLACK STUD. 936, 937 (2007); Linda van den Bergh et al., *The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap*, 47 AM. EDUC. RES. J. 497, 500 (2010).

73. *See* Russell J. Skiba et al., *The Color of Discipline: Sources of Racial and Gender Disproportionality in School Punishment*, 34 URB. REV. 317, 334 (2002).

74. *See* Steve Knotek, *Bias in Problem Solving and the Social Process of Student Study Teams: A Qualitative Investigation*, 37 J. SPECIAL EDUC. 2, 12 (2003).

75. Garda, *supra* note 34, at 1079 (2005).

76. *Id.* at 1078.

77. *See* Hallinan et al., *supra* note 37, at 96. Moreover, studies show that even when schools employ a set of criteria in placement decisions (most often grades, test scores, teacher and counselor recommendations, parental preference, and student choice), nonacademic factors play a significant role in determining the ability group level to which a student is assigned. *Id.*; *see also* Paula Stern & Richard J. Shavelson, *Reading Teachers' Judgments, Plans, and Decision Making*, 37 READING TCHR. 280, 281 (1983). Random factors—such as students' social skills, physical attractiveness, and style of dress—affect teachers' evaluations of student ability. *See* Ansalone, *supra* note 70, at 127.

78. *See* Losen, *supra* note 58, at 525.

79. *See id.*

in providing access to high-ability programs and gifted education for their children.[80]

Ethnic and class segregation is not merely a result of ability grouping but was also one of the motivations for ability grouping through the years.[81] In the early days of comprehensive schooling, ability grouping was a means to separate lower-class and immigrant children—who were largely uneducated—from those of the educated gentry.[82] After *Brown v. Board of Education*,[83] ability grouping expanded dramatically, coming to represent a means of circumventing desegregation by substituting intra-school segregation for what had previously existed between schools.[84] Despite typically being justified by educators as a response to student heterogeneity, the practice was undergirded by beliefs about race and class, and politically defended by white, middle-class parents seeking to preserve their privilege.[85] Ability grouping is therefore a central player in the construction of class and race relations in education—less conspicuous, perhaps, than de jure segregation but just as malignant.

The de facto segregation caused by ability grouping did not go unnoticed, as it attracted public criticism and even received legal challenges.[86] As a result, the practice of ability grouping saw a temporary drop toward the end of the twentieth century.[87] However, ability grouping has been on the upsurge in schools all over the country since the 2000s.[88] Over 70 percent of fourth-grade teachers who participated in a 2009 survey reported they had grouped students by reading ability, up from 28 percent in 1998.[89] In math, over 60 percent of fourth-grade teachers grouped students by ability in 2011, up from 40 percent in 1996.[90]

Concerns about the effect of ability grouping on the achievement gap between white and minority students have not eased

---

80. *See id.*

81. *See* Frank Biafora & George Ansalone, *Perceptions and Attitudes of School Principals Towards School Tracking: Structural Considerations of Personal Beliefs*, 128 EDUCATION 588, 589–90 (2008).

82. *See id.*

83. Brown v. Bd. of Educ., 347 U.S. 483 (1954).

84. Losen, *supra* note 58, at 521.

85. *See* OAKES, *supra* note 3, at 286–87.

86. *See* discussion *infra* Part IV.A.

87. LOVELESS, *supra* note 1, at 17.

88. *See id.* at 16.

89. *Id.* at 16.

90. *Id.* at 17.

with the resurgence of ability grouping in the last decade.[91] The evidence indicates ability grouping still correlates with socioeconomic status, race, and ethnicity.[92]

Ability grouping therefore seems to aggravate educational inequality by disadvantaging children of racial and ethnic minorities as well as poor children. The injustice caused far exceeds the realm of education and deeply affects students' life prospects. As a result, a shadow of doubt falls on the desirability of ability grouping as well as its moral permissibility.[93] This Article does not take a stand on the permissibility (or desirability) of ability grouping in general. Ability grouping is becoming more widespread than ever, practiced routinely in education systems with no signs of decline.[94] Therefore, while possibly not addressing all the concerns, reducing biases in the ability grouping process is an important contribution to educational justice.[95]

III. DATA-DRIVEN ABILITY GROUPING

There are no easy ways to eliminate implicit bias in education, as in other contexts.[96] Still, technology may offer a ray of hope. Decision-making processes that do not rely solely on human evaluations may be able to reduce biases in these processes. Ability grouping may be one of the practices that can benefit from new technologies.

A. Educational Data-Driven Decision-Making

As in many other life spheres, today's core educational activities rely increasingly on technological tools, such as digital whiteboards, digital textbooks, educational applications, mobile

---

91. *See* Richard R. Verdugo, *The Heavens May Fall: School Dropouts, the Achievement Gap, and Statistical Bias*, 43 EDUC. & URB. SOC'Y 184, 186 (2011).

92. *See* Werblow et al., *supra* note 58, at 272.

93. Several scholars argue to this effect. *See, e.g.*, CAROL CORBETT BURRIS & DELIA T. GARRITY, DETRACKING FOR EXCELLENCE AND EQUITY 50–65 (2008); Jo Boaler, *How a Detracked Mathematics Approach Promoted Respect, Responsibility, and High Achievement*, 45 THEORY INTO PRAC. 40 (2006); Hamsa Venkatakrishnan & Dylan Wiliam, *Tracking and Mixed-Ability Grouping in Secondary School Mathematics Classrooms: A Case Study*, 29 BRIT. EDUC. RES. J. 189, 201–02 (2003).

94. *See* LOVELESS, *supra* note 1, at 16–17.

95. One could argue that improving ability grouping would have the effect of further securing and embedding the practice, and therefore would have a negative overall effect on justice. However, a successful legal challenge to ability grouping in general is extremely unlikely, so it is better to improve ability grouping somewhat, even if it is impossible to solve all its problems.

96. *See* Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465, 473 (2010).

devices, online assessments, learning management systems (LMSs), and social networks.[97]

Interactive digital educational tools, such as those mentioned above, generate immense amounts of granular information about students.[98] This data—often called "big data"[99]—includes not only consciously disclosed information, such as entries concerning grades, behavior, and attendance, but also metadata concerning the students' online activity. Moodle, for example, is a popular LMS that can be used for task assignments, quizzes, content delivery, and communication.[100] Moodle logs students' every keystroke, including view and download commands, start and end time, time on task, and evaluation of assignments.[101]

In addition to the data collected from educational computerized platforms, further data concerning students may be accessible. Student ID cards may collect data on activities outside the classroom, such as purchases in the cafeteria or library loaning logs.[102] Schools may also collect information about students from email accounts, social media, and other noneducational sources.[103] Although not yet operational in most school systems, applications that can monitor bodily movements and indicators such as heart rate, eye movement, facial expressions, and posture already exist and can provide data concerning students' physical reactions while performing educational tasks.[104]

---

97. Most educators welcome the integration of technology to their classroom practices. *PBS Survey Finds Teachers Are Embracing Digital Resources to Propel Student Learning*, PUB. BROAD. SERV. (Feb. 3, 2013), http://www.pbs.org/about/blogs/news/pbs-survey-finds-teachers-are-embracing-digital-resources-to-propel-student-learning [https://perma.cc/M4YE-EE6X]. According to one survey, three-quarters of teachers expressed positive attitudes toward the integration of technology into the classroom. *Id.*

98. Elana Zeide, *The Limits of Education Purpose Limitations*, 71 U. MIAMI L. REV. 494, 505 (2017).

99. Big data is not easily defined, but in general refers to "large and complex datasets collected from digital and conventional sources that are not easily managed by traditional applications or processes." Jacqueleen A. Reyes, *The Skinny on Big Data in Education: Learning Analytics Simplified*, 59 TECHTRENDS 75, 75 (2015).

100. *See* Divna Krpan & Slavomir Stankov, *Educational Data Mining for Grouping Students in E-learning System*, PROC. 2012 34TH INT'L CONF. INFO. TECH. INTERFACES (ITI) 207, 208 (2012). Moodle is the acronym for Modular Object-Oriented Dynamic Learning Environment. *See id.* at 209.

101. *See id.*; Zeide, *supra* note 98, at 505.

102. Zeide, *supra* note 12, at 348–49.

103. *Id.* at 349.

104. Karen R. Effrem, *The Dark Side of Student Data Mining*, PULSE (June 3, 2016), http://thepulse2016.com/karen-r-effrem/2016/06/03/response-to-us-news-educational-data-mining-harms-privacy-without-evidence-of-effectiveness/ [https://perma.cc/HHC5-RTGH].

To make sense of the quantity and diversity of data, EDM technologies are used. EDM takes these seemingly unrelated data and finds unexpected correlations and patterns within them.[105] The connections between students' attributes, habits, and attainment offer opportunities for improving teaching and designing education policy: they can identify which students need help, and of which kind; they can inform educators about learning processes, what supports them, and what inhibits them;[106] and they help to evaluate teachers, courses, and pedagogical methods.[107] They can also inform educational policy, enabling multidimensional analysis at a level of detail and complexity previously unimaginable.[108]

One of the most dominant uses of EDM concerns assessments of students, teachers, schools, and school districts.[109] The use of information technologies for this purpose has largely been driven by legal requirements for data-based assessments and accountability.[110] Specifically, the No Child Left Behind Act (NCLB) imposes financial and administrative sanctions based on student test scores and focuses on closing the achievement gap in each school based on its

105. *See* Ryan S.J.D. Baker & George Siemens, *Educational Data Mining and Learning Analytics*, *in* THE CAMBRIDGE HANDBOOK OF THE LEARNING SCIENCES 253, 253 (R. Keith Sawyer ed., 2d ed. 2014). EDM refers to techniques, tools, and research designed to automatically extract meaning from large repositories of data generated by or related to people's learning activities in educational environments. R.S.J.D. Baker, *Data Mining for Education*, *in* INTERNATIONAL ENCYCLOPEDIA OF EDUCATION 112, 112–18 (B. McGaw et al. eds., 3d ed. 2010); *see also* Paul Baepler & Cynthia James Murdoch, *Academic Analytics and Data Mining in Higher Education*, 4 INT'L J. FOR SCHOLARSHIP TEACHING & LEARNING 1, 2 (2010); Félix Castro et al., *Applying Data Mining Techniques to e-Learning Problems*, 62 STUD. COMPUTATIONAL INTELLIGENCE 183, 184–85 (2007).

106. *See* VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, LEARNING WITH BIG DATA: THE FUTURE OF EDUCATION, 2 (2014).

107. Zeide, *supra* note 12, at 351.

108. *See* CTR. FOR DIG. EDUC., BIG DATA IN EDUCATION: HARNESSING DATA FOR BETTER EDUCATIONAL OUTCOMES 2 (2015); FED. TRADE COMM'N, BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION? 2 (2016); B.R. Prakash, M. Hanumanthappa & Vasantha Kavitha, *Big Data in Educational Data Mining and Learning Analytics*, 2 INT'L J. INNOVATIVE RES. COMPUTER & COMM. ENGINEERING 7515, 7516 (2014) (detailing the different kinds of insights EDM may offer).

109. Romero & Ventura, *supra* note 8, at 1–9.

110. *See* JULIE A. MARSH, JOHN F. PANE & LAURA S. HAMILTON, RAND CORP., MAKING SENSE OF DATA-DRIVEN DECISION MAKING IN EDUCATION: EVIDENCE FROM RECENT RAND RESEARCH 2 (2006), https://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf [https://perma.cc/C3QF-VV5B]. Legally established expectations for informed decision-making in education are not new and can be found in standards from as early as the 1980s and 1990s, which required the use of outcome data in school improvement planning and strategic planning. *See* ANDY HARGREAVES & HENRY BRAUN, NAT'L EDUC. POLICY CTR., DATA-DRIVEN IMPROVEMENT AND ACCOUNTABILITY 1–6 (2013), http://nepc.colorado.edu/publication/data-driven-improvement-accountability [https://perma.cc/R2NT-L629].

demographics and achievement scores.[111] To attain this goal, NCLB requires that states measure students' achievements annually and evaluate these achievements in light of state-established interim achievement goals, thereby making test scores and measurable student performance a primary concern for educators.[112] Race to the Top (RTT) also emphasizes accountability and measurement, while turning the focus from student achievement to student growth,[113] and offers states a considerable financial incentive to implement data-use policies and to invest in data-use infrastructure.[114] Despite slight differences between the two, both reforms drive the incorporation of data-rich technologies and EDM in schools.[115] In December 2015, new federal legislation was enacted: the Every Student Succeeds Act (ESSA).[116] This legislation is consistent with its predecessors, NCLB and RTT, in encouraging the use of accurate and transparent data on student performance.[117]

In addition to the ESSA, some states have also adopted policies to encourage the use of data for informing teachers' evaluations[118] and

---

111. No Child Left Behind Act of 2001 (NCLB), Pub. L. No. 107-110, 115 Stat. 1425 (2002) (repealed 2015); *see also* MARSH, PANE & HAMILTON, *supra* note 110, at 2.

112. NCLB § 1111(b)(2)(H). The state determines annually whether each district and school has made "Adequate Yearly Progress" (AYP). *Id.*; *see also* Robert L. Linn et al., *Accountability Systems: Implications of Requirements of the No Child Left Behind Act of 2001*, 31 EDUC. RESEARCHER 3, 3–4 (2002). A school does not meet the AYP if each subgroup of students does not improve in their proficiency levels. *See* NCLB § 1111(b)(2)(I); *see also* Linn et al., *supra*, at 3–4. Failing to meet the AYP entails sanctions on the school's and district's operation and autonomy. *See* Linn et al., *supra*, at 14.

113. DAMIAN W. BETEBENNER & ROBERT L. LINN, EDUC. TESTING SERV., GROWTH IN STUDENT ACHIEVEMENT: ISSUES OF MEASUREMENT, LONGITUDINAL DATA ANALYSIS, AND ACCOUNTABILITY 10 (2010), http://www.ets.org/Media/Research/pdf/BetebennerandLinnPresenterSession1.pdf [https://perma.cc/R2BQ-HZ7Y].

114. *See* HARGREAVES & BRAUN, *supra* note 110, at 3; Geoffrey H. Fletcher, *Race to the Top: No District Left Behind*, 37 TECH. HORIZONS EDUC. J. 17, 17–18 (2010); *see also* MEANS ET AL., *supra* note 9.

115. *See* HARGREAVES & BRAUN, *supra* note 110, at 1–6.

116. Every Student Succeeds Act (ESSA), Pub. L. No. 114-95, 129 Stat. 1802 (2015) (codified as amended at 20 U.S.C. §§ 6301–7981 (2012)). ESSA signifies a fundamental shift in terms of the relations between the federal government and the states by granting states more flexibility on issues related to accountability, resource allocation, and teacher evaluation. *See* AM. FED'N OF TEACHERS, EVERY STUDENT SUCCEEDS ACT: A NEW DAY IN PUBLIC EDUCATION, https://www.aft.org/sites/default/files/essa_faq.pdf [https://perma.cc/8KLC-XP5C] (last visited Sept. 3, 2017). States will be responsible for establishing their own accountability systems, though these must be submitted to and approved by the US Department of Education. *See* Paige Kowalski, *The Every Student Succeeds Act Says, "YES, Data Matter!"*, DATA QUALITY CAMPAIGN (Dec. 15, 2015), http://dataqualitycampaign.org/every-student-succeeds-act-says-yes-data-matter/ [https://perma.cc/GR6J-5G68]; AM. FED'N OF TEACHERS, *supra*.

117. *See* Kowalski, *supra* note 116.

118. *See* Clarin Collins & Audrey Amrein-Beardsley, *Putting Growth and Value-Added Models on the Map: A National Overview*, 116 TCHRS. C. REC. 1, 4 (2014) ("[Thirty] states and

instruction-related decisions.[119] To match the demand, a thriving industry of assessment systems has made these technologies readily available to teachers and schools.[120]

In addition to assessment and accountability driven by legislation, the data and data mining technologies are also used by schools for micro-decision-making,[121] such as ability grouping.[122]

B. Can Data-Driven Ability Grouping Reduce Biases?

In light of the persistent biases that plague traditional methods of educational decision-making, DDDM, with its purported scientific and objective nature, may make a welcome change. Data, it is argued, "doesn't lie";[123] therefore, decisions based on data mining results may be more objective and accurate than educators' judgment.[124] If, as research suggests, individuals are subconsciously prejudiced and evaluate identical data differently according to the relevant individual's race, social class, and sex,[125] machine-generated decisions may be preferable.

D.C. . . . now have legislation or regulations that require student achievement data be used to 'significantly' inform the criteria for the evaluation of teacher effectiveness . . . .").

119. *See* Deven Carlson, Geoffrey D. Borman & Michelle Robinson, *A Multistate District-Level Cluster Randomized Trial of the Impact of Data-Driven Reform on Reading and Mathematics Achievement*, 33 EDUC. EVALUATION & POL'Y ANALYSIS 378, 379–80 (2011).

120. *See id.* at 378–79.

121. The literature often characterizes DDDM in the educational context as a practice in which data is systematically collected, interpreted, and used for formulating action plans. Ellen B. Mandinach, Dir., Data for Decisions Initiative, WestEd, A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice, Address Before the 118th Annual Convention of the American Psychological Association (Aug. 2010), *in* 47 EDUC. PSYCHOLOGIST 71, 71 (2012). These action plans are continuously evaluated and adjusted based on further data. *See* Cynthia E. Coburn & Erica O. Turner, *The Practice of Data Use: An Introduction*, 118 AM. J. EDUC. 99, 104–05 (2012). This assumes that decision makers (educators, policy makers) have access to the data and are able to make sense of it, evaluate it, and then make informed decisions based on it. *See* Ellen B. Mandinach & Edith S. Gummer, *A Systemic View of Implementing Data Literacy in Educator Preparation*, 42 EDUC. RESEARCHER 30, 30–34 (2013).

122. *See* Vukicevic et al., *supra* note 8, at 146.

123. Arne Duncan, U.S. Sec'y of Educ., Address at the Fourth Annual IES Research Conference: Robust Data Gives Us the Roadmap to Reform (June 8, 2009), https://www.ed.gov/news/speeches/robust-data-gives-us-roadmap-reform [https://perma.cc/B42L-DF5F].

124. *See* Jeffrey R. Henig, *The Politics of Data Use*, 114 TCHRS. C. REC. 1 (2012); Mandinach, *supra* note 121, at 71. The US Department of Education promotes the collection and analysis of information generated by and about students as a means to help close achievement gaps, increase educational opportunities and college access, and reduce discrimination against underserved students. *See* MEANS ET AL., *supra* note 9, at 23, 25, 27.

125. *See supra* notes 69–72 and accompanying text.

Since the use of big data in education is in its early days, the evidence is still not conclusive as to its effect on biases in decision-making. However, initial evidence regarding DDAG suggests there is room for optimism.

EVAAS—an algorithm-based learning platform—provides data analysis services for the assessment of schooling effectiveness at the district, school, and classroom level by using various sources of information, including scores on standardized tests.[126]

EVAAS generates a multitude of assessments and predictions on teacher effectiveness, student proficiency, probability of success, risk of dropping out, and more.[127] According to the company's website, EVAAS is widely used to assign students to eighth-grade algebra.[128] The system evaluates a student's prior achievements to predict his or her success in higher-level courses and accordingly produces recommendations for assigning students to ability-based groups.[129]

Although systems such as EVAAS have not long been operational, research on their effect is already beginning to emerge. One study found that 19 percent of teachers who used EVAAS data stated that they used it for ability grouping, to differentiate instruction according to student ability, and to provide remedial education to those who needed it.[130] EVAAS's "probability of success" reports have also become a determinant factor in math placement policy in at least one school district.[131] Wake County in North Carolina decided achieving a certain level of success probability on

---

126. *See* S. PAUL WRIGHT ET AL., SAS INST. INC., WHITE PAPER: SAS® EVAAS® STATISTICAL MODELS 1 (2010), http://stat.wvu.edu/~wadillinger/Pres%202/EVAAS.pdf [https://perma.cc/H7F9-52AA]. For example, EVAAS uses testing scores provided by major educational testing companies and those used by states to fulfill their NCLB obligations. *See* SAS, *supra* note 13, at 2. On the other hand, EVAAS does not have access to students' social media activity, emails, and other online activities that are not school related. EDM, which has access to these types of data, may improve predictability even more and offer further insights into what makes students succeed. However, the ethical challenges that pertain to DDDM may also be more acute when these sources of information are included. *See* Xin Chen, Mihaela Vorvoreanu & Krishna Madhavan, *Mining Social Media Data for Understanding Students' Learning Experiences*, 7 IEEE TRANSACTIONS LEARNING TECH. 246, 246 (2014).

127. *See* WRIGHT ET AL., *supra* note 126, at 8, 10; SAS, *supra* note 13, at 10.

128. *See Expanding Eighth-Grade Algebra Participation*, *supra* note 13.

129. *See id.*

130. Clarin Collins, *Houston, We Have a Problem: Teachers Find No Value in the SAS Education Value-Added Assessment System (EVAAS®)*, 22 EDUC. POL'Y ANALYSIS ARCHIVES 1, 14 (2014).

131. *See* WAKE CTY. PUB. SCH. SYS., MIDDLE SCHOOL MATH PLACEMENT GUIDELINES, 2017–18, http://www.wcpss.net/cms/lib/NC01911451/Centricity/Domain/4039/Math%20Placement%20Guidelines%202017.18.pdf [https://perma.cc/4NPG-YZGJ] (last visited Sept. 3, 2017).

EVAAS's scale would be the criterion for assigning students to an accelerated track in math.[132]

Using EVAAS in assignment decisions instead of teacher recommendations increased the rates of African American, Latino, and low-income students in math acceleration.[133] The district also achieved proportional enrollment of female students: their enrollment in advanced math courses reflected their proportion in the student population.[134] Importantly, the measured success rates were not impacted by the change.[135]

An interesting finding concerns the reaction teachers had to the assignment recommendation EVAAS generated: when confronted with the assignment recommendations that EVAAS generated, teachers expressed surprise and admitted the model identified many students as suitable for the advanced course who otherwise would not have been chosen.[136]

Naturally, further research is required to investigate the variance between traditional methods of ability grouping and DDAG. Still, these initial findings are encouraging and suggest DDAG may offer opportunities for reducing biases and promoting equal educational opportunity.

That said, the use of data in itself "is not a panacea" for all ailments of educational inequality and may in fact create a new set of challenges in terms of equality.[137] Research into predictive analytics and data mining in other areas suggests that instead of eliminating biases, DDDM may reproduce them.[138] For example, algorithms used by the IRS to detect tax evaders, by police to detect potential drug offenders, and by banks to predict debtors who will be unable to repay their debt, have all been shown to produce predictions biased against racial minorities and people of lower socioeconomic status.[139]

---

132. *Id.*

133. *See* Dougherty et al., *supra* note 13, at 87S.

134. *Id.* at 87–89S.

135. *Id.*

136. *See Expanding Eighth-Grade Algebra Participation*, *supra* note 13.

137. Barocas & Selbst, *supra* note 15, at 673; *see* Cynthia E. Coburn & Erica O. Turner, *Research on Data Use: A Framework and Analysis*, 9 MEASUREMENT 173, 173 (2011).

138. *See, e.g.*, EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES 1, 47 (2014) [hereinafter PODESTA REPORT], https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf [https://perma.cc/H23L-Y9YB]. The so-called "Podesta Report" states that data mining may have unintended discriminatory effects: "The increasing use of algorithms to make eligibility decisions must be carefully monitored for potential discriminatory outcomes for disadvantaged groups, even absent discriminatory intent." *Id.*

139. *See* Barocas & Selbst, *supra* note 15; Kimberly A. Houser & Debra Sanders, *The Use of Big Data Analytics by the IRS: Efficient Solutions or the End of Privacy as We Know It?*, 19

Unequal outcomes in data-driven decisions are caused by preexisting social inequality that is merely reflected in the algorithm's output and by biases within the decision-making process, as is the case in traditional decision-making. This Article refers only to the latter and details the different ways in which this bias is created.[140]

1. Discriminatory Attributes

Algorithms learn how to make their predictions based on historical datasets.[141] To predict student success in a course, for example, algorithms analyze the data of past students (called the "training dataset") and find which attributes (or the complicated combination thereof) best predict student success.[142] If, historically, successful participants in honors classes have been mostly white and affluent, then the algorithm will try to locate similar candidates and inequality will be perpetuated.[143] Thus, biased decisions made in the past, as well as historical social inequality, are captured in the training dataset and resurface in the algorithms' predictions.[144]

To prevent this from happening, some algorithm scientists suggest removing discriminatory classifications such as race, gender, or ethnicity from the datasets.[145] If the algorithm does not have access to the racial identity of students, presumably it will not generate racially biased decisions.

2. Attributes That Correlate with Discriminatory Classifications

The problem with removing classifications such as race or sex from datasets is that other pieces of information that remain in the

---

VAND. J. ENT. & TECH. L. 817, 850 (2017); William Isaac & Andi Dixon, *Why Big-Data Analysis of Police Activity Is Inherently Biased*, SALON (May 13, 2017, 4:00 PM), https://www.salon.com/2017/05/13/why-big-data-analysis-of-police-activity-is-inherently-biased_partner/ [https://perma.cc/7ZFD-JLNH]; Katherine Noyes, *Will Big Data Help End Discrimination—Or Make It Worse?*, FORTUNE (Jan. 15, 2015), http://fortune.com/2015/01/15/will-big-data-help-end-discrimination-or-make-it-worse/, [https://perma.cc/P5FF-H9AZ].

140. As stated above, this Article does not deal with the ways in which law can address the background inequality that affects the achievement gap. In general, biases in the process of DDAG can be caused by problems in the data that algorithms analyze or by problems in the design of the algorithm itself. *See* PODESTA REPORT, *supra* note 138, at 6–10.

141. *See* Hajian & Domingo-Ferrer, *supra* note 32, at 242.

142. *See* Romero & Ventura, *supra* note 8, at 7.

143. *Cf.* PODESTA REPORT, *supra* note 138, at 8.

144. KAMIRAN & ŽLIOBAITE, *supra* note 17; Barocas & Selbst, *supra* note 15 at 671; Calders & Žliobaite, *supra* note 30, at 4.

145. Hajian & Domingo-Ferrer, *supra* note 32, at 241; Verwer & Calders, *supra* note 27, at 262.

data correlate with the discriminatory attributes.[146] For example, where residential segregation is severe, zip codes serve as a proxy for race and thus reintroduce racial bias into the algorithm's outputs.[147]

Removing all attributes that correlate with suspicious classifications could prove quite challenging because the correlation often stems from a combination of multiple types of data, such as activity in social media, online shopping habits, and interest or disinterest in specific online content.[148] Algorithms recognize these patterns and can obtain an accurate indication as to the individual's sex or race, even when the suspicious attributes (and those correlating with them) are removed.[149]

In addition to it being almost impossible to erase all traces of suspicious classifications from big datasets, removing these attributes can also be undesirable for other reasons.

First, removing certain attributes may decrease the accuracy of the algorithmic predictions.[150] This is the case when attributes that correlate with discriminatory classifications are relevant to educational decision-making. For example, the classification of students as English language learners (ELLs) correlates with immigrant status. Data on ELL eligibility may have to be excluded if immigration status is a classification we wish to remove from the database. This, however, is relevant data that could be important for optimal educational decision-making. Discipline and attendance reports may also correlate with suspicious classifications, yet they too seem like relevant inputs for optimal educational decision-making.[151]

An additional reason not to remove suspicious classifications from datasets is that the data collected can also be used for detecting educational inequality and for a deeper understanding of the mechanisms that create it. Removing these attributes makes it harder to monitor and contend with inequality.[152]

---

146. *See* Barocas & Selbst, *supra* note 15, at 712.

147. *Id.*; Verwer & Calders, *supra* note 27, at 262 (using the example of male-female and high income-low income).

148. *See* Barocas & Selbst, *supra* note 15, at 712.

149. *See id.*

150. Verwer & Calders, *supra* note 27, at 263.

151. *See* Skiba et al., *supra* note 73, at 333–34 (showing discipline is likely to correlate with race because there is inequality in the application of disciplinary policy with regard to African American students); *see also, e.g.*, OFFICE FOR CIVIL RIGHTS, U.S. DEP'T OF EDUC., CIVIL RIGHTS DATA COLLECTION: DATA SNAPSHOT: SCHOOL DISCIPLINE, CIVIL RIGHTS DATA COLLECTION 1 (2014), https://ocrdata.ed.gov/downloads/crdc-school-discipline-snapshot.pdf [https://perma.cc/X7JD-D36F] (finding that across all age groups, African American students were suspended and expelled at a rate three times greater than white students).

152. *See* Verwer & Calders, *supra* note 27, at 263.

3. Representation Within Data

Another challenge concerns the way members of protected classes are represented in the data.[153] A gap in technological proficiency separates students of privileged backgrounds—who commonly have high-quality Internet access at home—from less fortunate students.[154] Students who are less technologically proficient devote more time and cognitive resources to typing and navigating digital menus than to organizing and communicating ideas.[155] Studies have also found students of low-income families did not engage in online learning resources, and those who did, did not perform as well as their peers.[156] Even though the "digital divide"—the gap between high-income and low-income families in Internet access—is narrower than ever,[157] members of disadvantaged groups still lack the skills required to fully benefit from online educational resources.[158]

Finally, and more generally, the data available to algorithms are, necessarily, merely a reductive representation of an infinitely more specific real-world object or phenomenon. These representations may fail to capture the intricacies of reality.[159] Obtaining information rich enough to permit precise distinctions can be expensive, so data harvested as a side effect of existing activities are preferred. For example, data concerning the amount of time students are logged into

---

153. PODESTA REPORT, *supra* note 138, at 7–8.

154. *See* ELANA ZEIDE, FUTURE OF PRIVACY FORUM, 19 TIMES DATA ANALYSIS EMPOWERED STUDENTS AND SCHOOLS: WHICH STUDENTS SUCCEED AND WHY? 11 (2016), https://fpf.org/wp-content/uploads/2016/03/Final_19Times-Data_Mar2016-1.pdf [https://perma.cc/5F5L-ME3P] (finding that minorities, students of low socio-economic status, or ELLs are likely to have limited access to computers and Internet at home and therefore will be disadvantaged in a technology-based learning environment).

155. SHEIDA WHITE ET AL., NAT'L CTR. FOR EDUC. STATISTICS, PERFORMANCE OF FOURTH-GRADE STUDENTS IN THE 2012 NAEP COMPUTER-BASED WRITING PILOT ASSESSMENT: SCORES, TEXT LENGTH, AND USE OF EDITING TOOLS 63 (2015), https://nces.ed.gov/nationsreportcard/subject/writing/pdf/2015119.pdf [https://perma.cc/WP34-PECX]; *see also* ZEIDE, *supra* note 154, at 11.

156. Kaveh Waddell, *Virtual Classrooms Can Be as Unequal as Real Ones*, ATLANTIC (Sept. 26, 2016), https://www.theatlantic.com/technology/archive/2016/09/inequaity-in-the-virtual-classroom/501311/ [https://perma.cc/B9ZM-HCW9].

157. Andrew Perrin & Maeve Duggan, *Americans' Internet Access: 2000-2015*, PEW RES. CTR. (June 26, 2015), http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/ [https://perma.cc/7NQR-XME2] (finding that, according to census data, 84 percent of Americans now have Internet access and that for Americans aged 18–29, that figure is 96 percent); *see also* Monica Anderson & Andrew Perrin, *13% of Americans Don't Use the Internet. Who Are They?*, PEW RES. CTR.: FACT TANK (Sept. 7, 2016), http://www.pewresearch.org/fact-tank/2016/09/07/some-americans-dont-use-the-internet-who-are-they/ [https://perma.cc/GG4U-AU5D].

158. *See* Waddell, *supra* note 156.

159. *See* Calders & Žliobaite, *supra* note 30, at 47.

a LMS can be harvested at no cost, and, therefore, designers of algorithms often assign that considerable weight when deciding which students are likely to succeed (they assume that students who spend more time logged on are more likely to succeed in the course).[160] These data, however, do not necessarily communicate the whole story about the students' academic abilities and learning habits[161] and may be biased against students from poor backgrounds who tend to spend less time at home logged into the LMS.

The problems detailed above—concerning the data and the limited way the data represent reality—give rise to the possibility that DDAG may create new classes of individuals who are systematically educationally disadvantaged. These classes will include groups which, for some reason, are not properly represented in the data that are available to the algorithm, such as children who participate in after-school sports, or others. Given that educational disadvantage affects an individual's life prospects, this concern may prove significant.

4. Biases in the Design of the Algorithm Itself

Despite the fact that algorithms operate "independently" to discover connections that are simply "there" in the data, they are still—ultimately—designed and programmed by humans. Human biases can therefore seep into the process of data mining through the actions and decisions of the designers who program the algorithms.[162] Human involvement in algorithm design occurs at all stages: defining the attributes in the datasets, organizing the training datasets (functions referred to in the previous section), and determining the "question" the algorithm aims to answer.[163] This framing function is far from neutral. An algorithm used to assign students to a course, for

---

160. *See e.g.*, Angela Bovo et al., *Analysis of Students Clustering Results Based on Moodle Log Data*, 6TH INT'L CONF. ON EDUC. DATA MINING 306, 306 (2013); Krpan & Stankov, *supra* note 100.

161. Though some students may indeed spend this time learning, others may simply keep the window open while surfing the web or engaging in an online chat.

162. PODESTA REPORT, *supra* note 138, at 8–10; Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1517–20 (2013).

163. This model of data mining is called classification—a predictive data mining task. *See* Pedro G. Espejo, Sebastián Ventura & Francisco Herrera, *A Survey on the Application of Genetic Programming to Classification*, 40 IEEE TRANSACTIONS SYS. MAN & CYBERNETICS 121, 121 (2010). In other words, it aims to find connections among different attributes in the data that can best predict one specified attribute—success in a course, for example. *See id.* To make this prediction, the algorithm uses all the information it is fed, generates very high predictability rates, and finds surprising correlations between attributes that would not be established otherwise. *See id.*

example, can be programmed in various different ways: it can be asked to predict which students are most likely to succeed, it can identify the students with the highest ability, or it can be designed to determine which students are likely to benefit the most from the course. The different framing entails different assignment decisions and is therefore value-laden.

5. Why Are Biases Especially Troubling in Data-Driven Ability Grouping?

DDAG is, therefore, also susceptible to biases. In a certain respect, biases in DDAG are actually *worse* than biases in traditional ability grouping. The purported objectivity of algorithmic decision-making masks discrimination and prevents meaningful debate and critique.[164] As a result, discriminatory outcomes are excused and appear benign.[165]

This is especially problematic in education because, unlike other fields, the algorithms' predictions cannot be effectively verified ex post. After identifying potential tax evaders, an algorithm-based alert can be validated by an actual audit, and false predictions can be detected and corrected.[166] An innocent individual may be inconvenienced by being targeted by the algorithm, but this harm is relatively contained. Algorithms adjust as a result of these mistakes and improve their predictions. Conversely, a prediction that leads to the assignment of a student to a certain track does more than indicate the student's ability: it constitutes it. Teachers made aware of students' abilities unintentionally treat them differently in a manner that reinforces their perceptions of students' abilities.[167] Additionally, as ability grouping most often involves studying different curricula and allocation of unequal resources, students perceived as having higher ability are also granted better resources and taught superior skills, which further enhances their abilities. Disentangling the cumulative effects of the components of educational outcomes—prior ability, teacher expectations, differential resources, and curriculum—is therefore well nigh impossible. This hinders the ability to effectively validate the algorithm's initial prediction, making

164. Jules Polonetsky & Omer Tene, *Who Is Reading Whom Now: Privacy in Education from Books to MOOCs*, 17 VAND. J. ENT. & TECH. L. 927, 984–85 (2015); Polonetsky & Tene, *supra* note 12, at 31–32.

165. *Id.*

166. *See* Houser & Sanders, *supra* note 139, at 846–47.

167. *See* Lee Jussim, Stephanie Madon & Celina Chatman, *Teacher Expectations and Student Achievement: Self-Fulfilling Prophecies, Biases, and Accuracy*, *in* APPLICATIONS OF HEURISTICS AND BIASES TO SOCIAL ISSUES 303, 322 (Linda Heath et al. eds., 1994).

its outcomes virtually immune to critique.[168] It also significantly raises the stakes of the algorithms' decisions.

6. Possible Technological Solutions

In addition to removing suspicious classifications from the datasets, a move that the Authors do not find promising, scientists have begun devising technological solutions meant to contend with the biases that algorithmic decision-making may be prone to.[169]

One possibility involves the manipulation of training datasets to neutralize embedded biases. This activity in the service of equality involves choosing borderline cases concerning protected groups and changing their classification.[170] Thus, members of racial minorities who were not identified as suitable for higher tracks, but were close, would be reclassified as suitable. As a result, the algorithm would classify more members of racial minorities as suitable for higher tracks.

A more direct approach to creating an equal outcome could also be adopted. Algorithms can be programmed to produce equal outcomes, such as ability groups that fully reflect the population in terms of race, gender, or class. This would most likely entail modifying the decision threshold (for instance, average test scores), defining a different threshold of perceived ability for different ethnic or socioeconomic classes.[171] Doing so would immediately change the rate of children from racial minorities or low-income families assigned to higher tracks. This would also inevitably mean allocating fewer seats in higher tracks for students from privileged backgrounds (assuming that seats are limited).

Technologically, the problem with these two approaches (manipulating training datasets and producing predetermined equal outcomes) is they may decrease the algorithm's predictive accuracy. Assuming at least some of the inequality represented in the historical dataset or in current decisions results from actual social inequality rather than biases in decision-making, the algorithm would have to

168. *See* Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement,* 69 FLA. L. REV. 181, 181 (2016) (discussing possible methods of verification as an alternative to measures promoting transparency in algorithms).

169. *See* Hajian & Domingo-Ferrer, *supra* note 32, at 242–43; Verwer & Calders, *supra* note 27, at 263–68.

170. Hajian & Domingo-Ferrer, *supra* note 32, at 247–51.

171. Verwer & Calders, *supra* note 27, at 263.

consider race as a criterion for assignment recommendations and to apply different rules to students of different races.[172]

Arguably, a small decrease in accuracy should be tolerated if it leads to an improvement in equality. However, assuming ability grouping has a pedagogical justification, nonnegligible decreases in accuracy would be countereffective: they would entail assigning students to tracks unsuited to their ability and that do not fulfill their educational needs.

These solutions' differential treatments of individuals according to race also raise significant legal challenges, which are addressed in Part IV.

Another possible technological solution involves developing completely novel ways to group students. Typically, students are grouped according to their perceived abilities as evaluated by previous attainment or tests.[173] But algorithms can also offer other possibilities for grouping students, such as clustering them according to attributes other than ability. Clustering is a descriptive data mining model that groups together students with similar attributes.[174] These similarities would typically include, among other factors, their grades, knowledge in a particular field, capabilities, and skillsets,[175] but could also include more surprising categorizations such as learning styles,[176] habits, hobbies, and the like.[177] While this method would need to be empirically tested for pedagogical effectiveness, it offers a novel approach to grouping that may have a positive effect in terms of social integration.

---

172. If, on the other hand, inequality is caused wholly by biases in the process of decision-making, then these practices may actually improve accuracy.

173. Drowatzky, *supra* note 26, at 45–47.

174. *See* Neha D. & B.M. Vidyavathi, *A Survey on Applications of Data Mining Using Clustering Techniques*, 126 INT'L J. COMPUTER APPLICATIONS 7, 7 (2015). "The clusters that are formed need to satisfy the following two principles: 1) Homogeneity: Elements of the same cluster are maximally close to each other. 2) Separation: Data elements in separate clusters are maximally far apart from each other." *Id.*

175. *Id.* at 9.

176. Ioannis Magnisalis et al., *Adaptive and Intelligent Systems for Collaborative Learning Support: A Review of the Field*, 4 IEEE TRANSACTIONS LEARNING TECH. 5, 8 (2011). *See generally* Sofiane Amara et al., *Using Students' Learning Style to Create Effective Learning Groups in MCSCL Environments*, 1ST NAT'L CONF. ON EMBEDDED & DISTRIBUTED SYSTEMS (2015) (discussing the different methods of learning about students' learning styles).

177. *See* Romero & Ventura, *supra* note 8, at 9; *see also* Vukicevic et al., *supra* note 8, at 189; Ashish Dutt et al., *Clustering Algorithms Applied in Educational Data Mining*, 5 INT'L J. INFO. & ELECTRONICS ENGINEERING 112, 113 (2015); Li Li, Xiangfeng Luo & Haiyan Chen, *Clustering Students for Group-Based Learning in Foreign Language Learning*, 9 INT'L J. COGNITIVE INFORMATICS & NAT. INTELLIGENCE 55, 56–57 (2015).

As big data mining develops generally, and in the educational domain specifically, further technological solutions may be developed that might contend with inequality created through data mining.

IV. LEGAL REGULATION OF DATA-DRIVEN ABILITY GROUPING

After understanding the promises and pitfalls of big data for ability grouping, this Article examines two possible ways in which law can be instrumental in ensuring DDAG reduces biases and promotes equality: challenging cases in which DDAG results in racially biased decisions and regulating the design and practice of DDAG. This Article argues that legal challenges to unequal outcomes of DDAG are unlikely to be successful and suggests that the second strategy, namely regulating the design and practice of DDAG, is more promising.

A. Challenging Data-Driven Ability Grouping

The first way in which law can be instrumental in contending with inequality is through launching legal challenges to specific decisions or policies. This option, however, is unlikely to prove effective in the case of DDAG. The segregatory effects of traditional ability grouping policies have been challenged in courts several times, and though successful in some cases (to be detailed shortly), courts have, as a rule, upheld practices of ability grouping.[178] The difficulty to prove intentional discrimination and the continued disagreement among education experts as to the desirability of ability grouping have made the courts reluctant to strike down ability grouping policy.[179] DDAG is even more likely to withstand judicial review since it makes proving intentional discrimination even harder and arguably improves the grouping process by reducing race and class biases.

The first and most publicly known case to deal with the discriminatory effect of ability grouping was the 1969 case of *Hobson v. Hansen*.[180] The case challenged an ability grouping policy in the District of Columbia in which students were assigned to one of several tracks—from "basic" to "honors"—based on intelligence, achievement, and aptitude test scores.[181] The policy resulted in blatant segregation in schools: the higher tracks served an overwhelming majority of

---

178. *See* Losen, *supra* note 58, at 527–35.

179. *See id.*

180. Hobson v. Hansen, 269 F. Supp. 401, 406 (D.D.C. 1967), *aff'd sub nom.* Smuck v. Hobson, 408 F.2d 175 (D.C. Cir. 1969).

181. *See id.* at 406–07.

white students, whereas African American students were assigned mostly to lower tracks.[182] The district court ruled that although ability grouping was not illegal per se, the District of Columbia program violated the Due Process Clause of the Fifth Amendment.[183] In thus deciding, the court stressed the plaintiffs had been the victims of racial segregation throughout their prior education, and therefore, the tests used to perform the grouping did not give an accurate estimation of their ability.[184] The court also found "that education in the lower tracks was so watered-down as to be more fairly described as 'warehousing,'"[185] and the program did not involve review of the initial assignment decisions.[186] Therefore, the use of ability grouping in *Hobson* could not be understood as a temporary measure meant to help students overcome the educational disadvantage they suffered through segregation.

In *Moses v. Washington Parish School Board*, a court was faced with ability grouping in a recently desegregated school district.[187] Here, the previously white school absorbed all students and continued a grouping system it practiced prior to desegregation, which comprised eleven homogeneous levels.[188] Tracking in principle was not held to be illegal in this case either; instead, the decision to strike down the policy was based on the fact that the students who studied in segregated schools had received inferior prior education.[189]

Despite these successes, the applicability of these precedents was critically limited in subsequent cases.[190] The *Hobson* court was clear that ability grouping is not unlawful per se[191] and that it is a legitimate education policy when it is reasonably related to a legitimate educational objective and implemented in a nonarbitrary, noncapricious, and nondiscriminatory way. The subsequent jurisprudence distinguished school districts operating under preexisting desegregation orders from those that had reached unitary

---

182. *See id.* at 456.

183. *Id.* at 511.

184. *Id.* at 514 ("[R]ather than being classified according to ability to learn, these students are in reality being classified . . . *according to environmental and psychological factors which have nothing to do with innate ability*." (emphasis added)).

185. *See* Smuck v. Hobson, 408 F.2d 175, 187 (D.C. Cir. 1969).

186. *See id.*

187. *See* Moses v. Washington Par. Sch. Bd., 330 F. Supp. 1340, 1340 (E.D. La. 1971), *aff'd*, 456 F.2d 1285 (5th Cir. 1972).

188. *Id.* at 1341.

189. *Id.* at 1345.

190. *See* Losen, *supra* note 58, at 529.

191. *See Smuck*, 408 F.2d at 186.

status or had never been under desegregation orders.[192] In school districts operating under a desegregation order, evidence of segregation in ability grouping raises a presumption of discriminatory intent and, therefore, the burden of proof shifts to the district to show that the policy is not a vestige of that original discrimination.[193] On the other hand, this presumption does not apply to districts operating under unitary status for sufficient time.[194]

In *NAACP v. Georgia*, the ability grouping practice involved students who had not attended segregated schools themselves, despite the fact that the district was under a desegregation order and had not achieved unitary status.[195] The court found that segregation could not be blamed for the inequality in educational abilities that was reflected in the racially disparate grouping outcomes.[196] The fact that the students' parents attended segregated schools and the school district still had not achieved unitary status was deemed irrelevant to the current grouping system.[197] More importantly, the court deferred to the district's opinion that ability grouping was a legitimate educational practice (including tracking students as early as kindergarten), and moreover, that ability grouping could offer remedial education for racial minorities.[198]

Since *NAACP*, challenges to practices of tracking based on racial imbalance have been tough battles to win without proof of

---

192. *See* Losen, *supra* note 58, at 530.

193. *See* Simmons *ex rel.* Simmons v. Hooks, 843 F. Supp. 1296, 1302 (E.D. Ark. 1994) ("Ability grouping which results in racial segregation may be permitted in an otherwise unitary school system if the school district can demonstrate that its ability grouping is not based on the present results of past segregation or that it will remedy such results through better education opportunities."). Cases involving ability grouping in school districts under desegregation orders also include objections to districts' motions seeking unitary status; courts sometimes grant unitary status despite the district's failure to satisfy all the requirements. *See, e.g.*, Freeman v. Pitts, 503 U.S. 467, 471 (1992).

194. McNeal v. Tate Cty. Sch. Dist., 508 F.2d 1017, 1020–21 (5th Cir. 1975). In *McNeal*, the US Court of Appeals for the Fifth Circuit upheld the prohibition of an ability grouping practice because the district failed to show that its student assignment methods were "*not* based on the present results of past segregation." *Id.* (emphasis added). However, this statement was not intended as a hard-and-fast rule, and the court did leave space for cases in which evidence might show that a given system of grouping is in students' best interests. *Id.*

195. Ga. State Conference of Branches of NAACP v. Georgia, 775 F.2d 1403, 1413–14 (11th Cir. 1985); *see also* Montgomery v. Starkville Mun. Separate Sch. Dist., 854 F.2d 127, 130 (5th Cir. 1988) (finding that past segregation could not be blamed for ability grouping's disparate impact because the school district had been under a desegregation order for twenty years).

196. *NAACP*, 775 F.2d at 1414, 1416.

197. *Id.* at 1414–15.

198. *Id.* at 1410, 1419.

intent to discriminate.[199] Courts have repeatedly upheld ability grouping policies despite the racial imbalance that ensued.[200] And while school districts that were under desegregation orders in the past are "considerably more vulnerable to equal protection arguments" than those that were not, willingness to intervene even in those cases is small.[201] Equal protection challenges therefore have become ineffective unless intentional discrimination can be proved.[202]

199. *See* Quarles v. Oxford Mun. Separate Sch. Dist., 868 F.2d 750, 753 (5th Cir. 1989) ("[A]bility grouping has been recognized by both courts and educators as an acceptable and commonly used instruction method"); *Montgomery*, 854 F.2d at 130 ("We are impressed particularly with the testimony . . . [that] achievement grouping is far superior to ability grouping."); *NAACP*, 775 F.2d at 1419 ("The district court's findings of the educational soundness of interclass ability arrangements *per se* are not clearly erroneous. The record discloses that such grouping permits more resources to be routed to lower achieving students in the form of lower pupil-teacher ratios and additional instructional materials."). *But see* United States v. Yonkers Bd. of Educ., 123 F. Supp. 2d 694, 718 (S.D.N.Y. 2000) (holding that since ability grouping in the district was based on teachers' "attitudes and expectations" that could be traced to prior segregation, the ability groups themselves were a form of segregation); Simmons *ex rel.* Simmons v. Hooks, 843 F. Supp. 1296, 1302 (E.D. Ark. 1994) (applying the *McNeal* test and finding that tracking could not remedy the results of past discrimination).

200. *See* People Who Care v. Rockford Bd. of Educ., 111 F.3d 528, 536 (7th Cir. 1997); Price v. Austin Indep. Sch. Dist., 945 F.2d 1307, 1313 (5th Cir. 1991) (ruling that once the school system has been held "unitary," the burden shifts to the plaintiff to show that a newly adopted student assignment plan with a disparate impact on minorities is intentionally discriminatory); *Quarles*, 868 F.2d at 754 (conceding that there was "a high concentration of white students in the upper level groups" and a high concentration "of Black students in the lower level groups," but holding that this was not a result of the school's former segregated system).

201. *See* Losen, *supra* note 58, at 532. The stronger protection offered in districts that were segregated does not apply to a growing number of racial and ethnic minority children whose ancestors did not attend segregated schools, either because they did not reside in southern states or because they immigrated to the United States after *Brown v. Board of Education*, 347 U.S. 483 (1954). *See, e.g.*, Castaneda v. Pickard, 648 F.2d 989, 998 (5th Cir. 1981). In *Castaneda*, the tests used for ability grouping in the Raymondville Independent School District (RISD) were administered entirely in English, so all ELLs were placed in the "low-ability" group. *Id.* The US District Court for the Southern District of Texas nonetheless ruled in favor of RISD. *Id.* at 1015. On appeal, the Fifth Circuit partially reversed on other grounds. *See id.*; *see also* Douglas S. Reed, Legal and Pedagogical Contexts of English Learners: Defining "Appropriate Action" Under the Equal Educational Opportunity Act 11 (Mar. 29, 2016) (unpublished manuscript) (on file with author).

202. Losen, *supra* note 58, at 529. Angelia Dickens suggests that the Court should adopt a belief in the "fundamentality of education" adopted by Justice Marshall in his dissent in *San Antonio Independent School District v. Rodriguez*, 411 U.S. 1, 116 (1973) (Marshall, J., dissenting), and further argues that the practice constitutes a classification based on race that should be subject to strict scrutiny. *See* Angelia Dickens, Note, *Revisiting* Brown v. Board of Education*: How Tracking Has Resegregated America's Public Schools*, 29 COLUM. J.L. & SOC. PROBS. 469, 485, 488 (1996). Thus, under Dickens's formulation, a school district would be required to show that ability grouping is "narrowly tailored to serve a compelling state interest." *See id.* at 500. In her view, a district will likely not be able to establish a compelling interest for tracking; therefore, an Equal Protection challenge to ability grouping under her framework for strict scrutiny analysis would likely succeed. *Id.*

Claims brought under Title VI of the Civil Rights Act of 1964[203] are also insufficient for challenging racial biases in ability grouping. Title VI does not require proof of intentional discrimination and can apply when ability grouping results in significant levels of classroom segregation.[204] However, policies causing an indirect disparate impact can be redeemed, according to Title VI, if they are justified from an educational perspective and are the least segregatory out of equally effective educational alternatives.[205] As previously noted, courts have deferred to professional expertise as to whether ability grouping is overall better for students[206] and have refrained from seriously considering the possibility that even good faith efforts at grouping could be biased.[207]

Courts' acceptance of ability grouping as a legitimate educational practice—even when it results in racial segregation—is a key barrier to legal challenges of the practice. The view taken by courts impedes Title VI claims and prevents bringing forward claims according to the rational basis test, which applies both to nonsuspicious classifications, such as socioeconomic class,[208] and when the rights that are being infringed upon are not "fundamental."[209] To successfully challenge a state action, plaintiffs are required to prove that it bears no rational relation to a legitimate governmental interest.[210] This would be nearly impossible to prove, considering

---

203. Title VI is a general antidiscrimination law that bars discrimination on the basis of race and national origin in programs and services operated by recipients of federal financial assistance. *See* Civil Rights Act of 1964, Pub. L. No. 88-352, § 601, 78 Stat. 252, 252–53 (codified as amended at 42 U.S.C. § 2000d (2012)). Ability grouping policies or processes that operate to discriminate on the basis of student gender are also prohibited by Title IX of the Education Amendments of 1972. *See* Education Amendments of 1972, Pub. L. No. 92-318, § 901, 86 Stat. 373, 373–75 (codified as amended at 20 U.S.C. §§ 1681–86 (2012)).

204. *See* 34 C.F.R § 100.3(b)(2) (2017). However, the efficacy of Title VI disparate impact claims has been questioned. *See, e.g.*, Olatunde C.A. Johnson, *Disparity Rules*, 107 COLUM. L. REV. 374, 396 (2007).

205. *See* 42 U.S.C. § 2000d (2012); Guardians Ass'n v. Civil Serv. Comm'n of the City of New York, 463 U.S. 582, 624 n.15 (1983).

206. *See* Montgomery v. Starkville Mun. Separate Sch. Dist., 854 F.2d 127, 130 (5th Cir. 1988). In the *NAACP* case, the court referred to both equal protection claims and claims under Title VI. Ga. State Conference of Branches of NAACP v. Georgia, 775 F.2d 1403, 1408 (11th Cir. 1985). The court ruled that a racially disparate grouping system did not violate Title VI because grouping was necessary to meet the needs of the student population and was an "accepted pedagogical practice." *Id.* at 1418 (quoting the district court record).

207. Note, *Teaching Inequality: The Problem of Public School Tracking*, 102 HARV. L. REV. 1318, 1326 (1989).

208. *See* San Antonio Indep. Sch. Dist. v. Rodriguez, 411 U.S. 1, 27–28 (1973).

209. *Id.* at 51.

210. *See* FCC v. Beach Commc'ns, Inc., 508 U.S. 307, 313 (1993). The Court's application of the rational basis test has made it so permissive that it is practically unusable. In one case, the Court explained that "if there is any reasonably conceivable state of facts that could provide

courts have repeatedly accepted ability grouping as a legitimate, and therefore rational, policy choice.

Existing equal protection jurisprudence, therefore, has been largely ineffective in safeguarding equality of opportunity for disadvantaged groups. As the United States moves away from the painful history of de jure segregation, the possibility of courts applying a stricter standard of review decreases even more. In the case of DDAG, the existing doctrines are even less likely to be effective in challenging the unequal effects of ability grouping. Algorithmic decision-making is perceived as scientific and objective; therefore, courts are even more likely to defer to the grouping decisions made by algorithms, which renders both Title VI and the rational basis test under the Due Process Clause ineffective.[211] Moreover, intentional discrimination can easily be disguised in algorithmic decision-making

---

a rational basis" for a challenged law, it will survive rational basis review. *Id.* Moreover, the Court stated that it was irrelevant whether the rationale given for the challenged distinction actually motivated the legislature, suggesting that any plausible reason can suffice, whether or not it was the true reason for legislation. *See id.*; *see also* Jeffrey D. Jackson, *Putting Rationality Back into the Rational Basis Test: Saving Substantive Due Process and Redeeming the Promise of the Ninth Amendment*, 45 U. RICH. L. REV. 491, 493 (2011). Additionally, the standard of proof required of plaintiffs is extremely high, creating a "virtually irrebuttable presumption of constitutionality under the rational basis test." Clark Neily, *No Such Thing: Litigating Under the Rational Basis Test*, 1 N.Y.U. J.L. & LIBERTY 898, 908 (2005). In short, the rational basis test is extremely unlikely to be helpful in addressing cases of racial bias in ability grouping. As Jackson, *supra*, at 493, notes, "the Court has essentially made the rational basis test the equivalent to no test at all." *But see* Robert C. Farrell, *Successful Rational Basis Claims in the Supreme Court from the 1971 Terms Through* Romer v. Evans, 32 IND. L. REV. 357 (1999) (counting ten cases in twenty-five years in which this rational basis "with a bite" has been applied, in contrast to the one hundred cases in which it has been rejected); Gerald Gunther, *Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1, 21 (1972) (noting that there have been several cases over the years in which courts have applied a more stringent version of the rational basis test); Gayle Lynn Pettinga, Note, *Rational Basis with Bite: Intermediate Scrutiny by Any Other Name*, 62 IND. L.J. 779 (1987).

211. *See* Barocas & Selbst, *supra* note 15, at 677. At first glance, EDM seems extremely successful in terms of the rational basis test, as it is a good predictor of educational success. There is, however, something special in algorithmic decision-making that raises doubt as to the appropriateness of the rational basis test as a matter of principle. The "point of data mining is to provide a rational basis upon which to distinguish between individuals and to reliably confer to the individual the qualities possessed by those who seem statistically similar." *Id.* The statistical correlations that algorithms find are always rational in the sense that they are statistically valid. Therefore, any finding of an algorithm is rational and passes the legal test. However, its inexplicable "black box" nature raises doubt as to whether decisions generated from it can satisfy the rational basis test. For a mechanism to be rational, it must offer some substantive explanation for its decisions. *See* Farrell, *supra* note 210, at 383. Another problem is that in each and every prediction offered by the algorithm, the explanation for why its predictions always supposedly satisfy the rational basis test would be "because the algorithm said so." Absent a possibility to sometimes fail the test, the rational basis test seems to have no meaning at all: when everything is rational, nothing is rational.

behind complicated correlations.[212] Therefore, attempts to utilize ability grouping to preserve racial and class segregation would be even harder to combat.

Another important barrier in placing challenges before DDAG is the lack of transparency of algorithms. Several scholars advocate for promoting due process rights in DDDM.[213] The key, it seems, is ensuring decision makers, as well as individuals affected by the decision, can review and challenge the decision. For this, transparency and interpretability are crucial. Transparency,[214] through code disclosure or otherwise, will enable educators to review the data and make decisions based on it without surrendering their discretion to machines. Transparency will also enable students to access their information, correct it, and know how they are rated.[215]

The problem with requiring transparency is that algorithms are extremely opaque, making disclosure only minimally helpful. Hence, a precursory requirement for fostering transparency is interpretability.[216] The outcome and the way it was reached should be simplified—perhaps through graphic display—so that students, parents, and teachers can understand it.[217] The complex processes are not only inaccessible in terms of human understanding but also are often legally protected trade secrets—blocking anything but very general descriptions of the processes leading to the predictions.[218] As a result, students affected by the algorithms' recommendations have limited ability to understand the rationale behind the decision and to challenge it.[219] Making the factors that are considered by algorithms

212. Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 682 (2017).

213. *See e.g.*, Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014); Zarsky, *supra* note 162, at 1547.

214. *See* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015).

215. *See id.*; Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 24 (2014).

216. *See* Philipp Hacker, *Nudge 2.0—The Future of Behavioural Analysis of Law, in Europe and Beyond: A Review of 'Nudge and the Law: A European Perspective'*, 24 EUR. REV. PRIV. L. 297, 308–09 (2016); Richard H. Thaler & Will Tucker, *Smarter Information, Smarter Consumers*, HARV. BUS. REV., Jan.–Feb. 2013.

217. *See* Julia Stoyanovich & Ellen P. Goodman, *Revealing Algorithmic Rankers*, FREEDOM TO TINKER (Aug. 5, 2016), https://freedom-to-tinker.com/2016/08/05/revealing-algorithmic-rankers/ [https://perma.cc/6QRR-WJT5] (arguing that transparency, wherein the rules of operation of an algorithm are more or less apparent, or even fully disclosed, still leaves stakeholders in the dark). Instead, Stoyanovich and Goodman advocate for interpretability "which rests on making explicit the interactions between the program and the data on which it acts." *Id.*

218. *See* PASQUALE, *supra* note 214, at 14–15.

219. *See* Barocas & Serbst, *supra* note 15, at 696.

publicly known might also allow for strategic behavior aimed at getting high scores.[220]

Finally, due process rights inevitably entail reintroducing human biases into the decision-making process. If teachers are able to override algorithms' recommendations and assign children who were not identified by the algorithm to a higher track, it would not be surprising if this discretion were practiced more often in favor of children from privileged families.

It is also likely that allowing students to appeal DDAG decisions would benefit children of privileged families, because they are typically better equipped to take advantage of due process rights than students from disadvantaged families.

The discussion above suggests that challenging specific assignment decisions using traditional doctrines of equal protection is unlikely to succeed in ensuring that DDAG will promote educational equality and decrease biases. Law may be more effective in ensuring these goals by being involved in the design and implementation of the algorithms used in ability grouping. To this end, this Article suggests integrating technological solutions and legal regulation.

B. Regulating the Design and Implementation of Data-Driven Ability Grouping

Challenging assignment decisions or ability grouping policies in courts is not a promising route for promoting equality. Instead, this Article argues that law can be more effective if it is involved in the design and application of DDAG. The development and design of algorithms that are sensitive to equality are in their first steps. As a result, this Article does not purport to offer any comprehensive solution here. Instead, it aims to describe what such future solutions may look like and offers some insights into the way technological and legal solutions ought to be integrated to achieve the ultimate goal.

Algorithms function as policies.[221] They determine criteria for allocating certain resources or entitlements which are then applied to individuals. They are much easier to regulate than human decision-making because once the criteria are set and the weight given to each attribute is assigned, the algorithm reliably follows its own rules. Although this does not prevent biases from infiltrating—as described at length above—it does mean that technological intervention to correct biases can be effective, as opposed to

220. *See* PASQUALE, *supra* note 214, at 217.

221. *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1254 (2008).

irreparably biased human decision-making. These characteristics make it possible to envisage DDAG as a means to promote equality in education.

Legal intervention is required in the design of equality-sensitive algorithms for two main reasons. First, designing equality-sensitive algorithms entails normative determinations that legal doctrine and scholarship are best equipped to make. Second, legal regulation ensures universal implementation. Creating the technological tools to decrease biases requires expertise and may be costly, so legal regulation is essential to ensure that all schools and school districts using DDDM implement bias-reducing systems.

Scientists are aware of the biases that may be perpetuated by EDM and have begun devising technological solutions.[222] These attempts are commendable because developing technological solutions can optimize DDAG and promote educational equality. But these solutions inherently involve a myriad of normative decisions that law needs to address: Which groups warrant special attention (race, gender, class)? What does an equal or fair outcome consist of—equal shares or something different? Is differential treatment acceptable?

For example, an algorithm may be designed to assign zero weight to race, arguably creating a race-neutral assignment mechanism. Conversely, algorithms can be designed to create equal racial representation, thus instating differential criteria for students of different racial groups. A third possibility involves manipulating the historical datasets and offsetting some of the existing bias. Each choice will result in different outcomes—in terms of both specific assignment decisions and in the level of segregation in the education system as a whole. The choice between the different options is not technological but normative. Each choice expresses a different understanding of what fair assignment policy requires.

Unfortunately, research on these issues in the computer science community has not had recourse to the highly sophisticated and developed legal doctrine and scholarship.[223] As a result, these efforts may fail to appropriately address the problems identified in DDAG. Technological solutions must meet the goals set by normative and legal dictates. Legal involvement is important not only to direct

---

222. *See supra* Part III.B.

223. For an attempt at integrating legal and technological perspectives in discovering discrimination, see Dino Pedreschi, Salvatore Ruggieri & Franco Turini, *The Discovery of Discrimination*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES, *supra* note 17, at 91. For a general overview of this area, also see the other chapters of DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES, *supra* note 17.

the design of algorithms but also to ensure that effective technological solutions are uniformly applied to all cases of DDAG.

To design algorithms that will reduce biases, we must consider a complicated set of empirical questions including, but not limited to, whether applying equal criteria to all children imposes differential burdens on children of diverse background; whether students have been exposed to prior injustice; what is the threshold of ability required for benefiting from a course; and what the side effects will be of each mode of assignment. The answers to these empirical questions are to be found within the expertise of educators and social scientists. The normative discussion must react to these facts, determining the normative commitments and the legal framework within which they can be realized.

Earlier, the Authors distinguished unequal outcomes caused by social inequality (which existed before the grouping decision and is unrelated to it) from those caused by biases in the decision-making process.[224] This distinction resurfaces now, when the Authors are required to decide whether to design algorithms merely to reduce biases within the grouping process or to engage in the more ambitious task of minimizing the reflection of social inequality in ability grouping.[225]

Designing algorithms to correct anything but biases in the decision-making process itself may reasonably be classified as affirmative action, which in the current legal atmosphere is a "nonstarter."[226] Courts have struck down policies that treat members of different racial groups differently even when this differential treatment was designed to facilitate integration and promote equal opportunity.[227] To withstand strict scrutiny, educational policy that gives preferential treatment to racial minorities must promote a compelling state interest and be sufficiently narrowly tailored.[228] In the seminal case of *Parents Involved*, the Supreme Court struck down assignment policies in two school districts that considered students' race in assigning them to schools, even though this policy's objective was to promote racial diversity.[229] In striking down the policy, the

224. *See supra* Part II.B, particularly notes 67–80 and accompanying text.

225. This involves cases when the algorithm "goes too 'right,'" as Barocas and Selbst put it, and social inequality is to blame for the unequal outcome. Barocas & Selbst, *supra* note 15, at 729. Achieving equal outcome despite social inequality would require, as noted above, decreasing the predictive accuracy of algorithms. *See supra* Part III.B.6.

226. *See* Barocas & Selbst, *supra* note 15, at 715.

227. *See* Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 742 (2007).

228. *See id.*

229. *See id.* at 710–11.

Court stated it was not sufficiently narrowly tailored,[230] and while race may be considered, it could only constitute one consideration among many—students must be evaluated holistically rather than merely according to their race.[231] Following *Parents Involved*, the US Department of Education Office for Civil Rights and the US Department of Justice Civil Rights Division issued joint Diversity Guidelines for school districts, in which they detail the measures school districts may adopt to promote diversity in a constitutional manner.[232] The guidelines advise school districts first to examine race-neutral measures and then use generalized race-based approaches that do not refer to any specific student.[233] Individualized racial examination should be used as a last resort and be narrowly tailored to the district's specific goals.[234] In these cases race may be considered alongside other considerations in assessing a student's assignment.[235] These guidelines do not refer explicitly to ability grouping, but the rationale seems to apply directly. They suggest that as long as race is merely one consideration among many others and students are evaluated holistically, school districts are allowed to consider it in order to realize the compelling state interest of racial integration.[236]

Since algorithms incorporate multiple considerations other than race, it seems that some of the means to promote racial equality in assignment may withstand strict scrutiny under *Parents Involved.*

Additionally, while the focus on race is understandable, it is important to keep in mind that racial disparities are not the only inequalities that ability grouping recreates. Children of lower socioeconomic class are also overrepresented in lower tracks, as are immigrants. Gender inequality is also an issue, especially in science, technology, engineering, and mathematics (STEM) courses. These classifications are not probed as strictly by courts, requiring only

---

230. *Id.* at 726. Four of the five majority justices went further to state that racial diversity was not a compelling state interest. *Id.* at 730–32. Justice Kennedy, however, joined the dissent in asserting that integration was a compelling state interest. *Id.* at 788–89 (Kennedy, J., concurring in part and concurring in the judgment). This case has been subject to wide scholarly critique. *See, e.g.*, Philip Tegeler, *The 'Compelling Government Interest' in School Diversity: Rebuilding the Case for an Affirmative Government Role*, 47 U. MICH. J.L. REFORM 1021 (2014).

231. *See Parents Involved*, 551 U.S. at 723.

232. *See* CIVIL RIGHTS DIV., U.S. DEP'T OF JUSTICE, & OFFICE FOR CIVIL RIGHTS, U.S. DEP'T OF EDUC., GUIDANCE ON THE VOLUNTARY USE OF RACE TO ACHIEVE DIVERSITY AND AVOID RACIAL ISOLATION IN ELEMENTARY AND SECONDARY SCHOOLS (2011), https://www2.ed.gov/about/offices/list/ocr/docs/guidance-ese-201111.pdf [https://perma.cc/W8LV-8YAJ].

233. *Id.* at 7.

234. *See id.*

235. *See id.*

236. *See id.*

intermediate scrutiny (in the case of gender and nationality) or the lenient rational basis test in the case of socioeconomic status.[237] As a result, it would seem algorithms designed to correct biases would withstand judicial review.

To conclude, DDAG is a case in which legal intervention can be most effective in the stage of design and policy making. To make the most of what DDAG has to offer, though, cooperation is needed among scientists, educators, and lawyers. Bridging this professional gap is the practical challenge currently confronting policy makers.

V. CONCLUSION

Brown marked the beginning of the end of de jure segregation in the United States. But segregation in education did not end; rather, it underwent modification. Attending the same school is hardly a remedy for school segregation if African Americans and whites are separated upon entering the schoolhouse doors. Regardless of the policy's alleged neutrality, minorities are disadvantaged by tracking when the assignment of students creates separate and racially identifiable classrooms, which, in turn, provides minorities with fewer educational resources and opportunities and inferior life prospects.

Technological developments, more specifically EDM, have the potential to improve the ability grouping process and to begin to deliver long-promised educational justice to all children. Whether DDAG will ultimately succeed depends on multiple factors, of which legal regulation is only one. Educators and regulators alike must watch the implementation of DDAG carefully and adjust its design as its effects become known. If, after all, DDAG is unable to promote equality of opportunity and decrease segregation—both between and within schools—there may be no choice but to revisit the struggle to eliminate ability grouping altogether.

---

237. *See, e.g.*, San Antonio Indep. Sch. Dist. v. Rodriguez, 411 U.S. 1, 18 (1973) (deciding that socioeconomic status was not a suspicious classification that triggers strict or intermediate scrutiny). Noting the difference in jurisprudence between categories of race and class, several writers suggest promoting equality and diversity by using socioeconomic class instead of race. *See* Eboni S. Nelson, *The Availability and Viability of Socioeconomic Integration Post-*Parents Involved, 59 S.C. L. REV. 841 (2008); Kimberly Jenkins Robinson, *The Constitutional Future of Race-Neutral Efforts to Achieve Diversity and Avoid Racial Isolation in Elementary and Secondary Schools*, 50 B.C. L. REV. 277 (2009); James E. Ryan, *The Supreme Court and Voluntary Integration*, 121 HARV. L. REV. 131 (2007); Ronald Turner, *The Voluntary School Integration Cases and the Contextual Equal Protection Clause*, 51 HOW. L.J. 251 (2008); Lauren E. Winters, *Colorblind Context: Redefining Race-Conscious Policies in Primary and Secondary Education*, 86 OR. L. REV. 679 (2007).