Georgia Educational Researcher

Volume 17 | Issue 1 Article 3

1-2020

Testing for All: An Exploration of Disproportionality in the Georgia Alternate Assessments

Kelly D. Childers
Valdosta State University, kellydale31@gmail.com

Meagan Arrastia-Chisholm
Valdosta State University, mcarrastia@valdosta.edu

Katharine S. Adams
Valdosta State University, ksadams@valdosta.edu

Heather M. Kelley Ph.D.
Valdosta State University, hkelley@valdosta.edu

Recommended Citation

Childers, Kelly D.; Arrastia-Chisholm, Meagan; Adams, Katharine S.; and Kelley, Heather M. Ph.D. (2020) "Testing for All: An Exploration of Disproportionality in the Georgia Alternate Assessments," *Georgia Educational Researcher*: Vol. 17 : Iss. 1 , Article 3.
DOI: 10.20429/ger.2020.170103
Available at: https://digitalcommons.georgiasouthern.edu/gerjournal/vol17/iss1/3

Testing for All: An Exploration of Disproportionality in the Georgia Alternate Assessments

Abstract

Public schools use standardized testing to measure students' academic achievement at the conclusion of each school year. Students with severe cognitive disabilities are evaluated through the Georgia Alternate Assessment (GAA). The purpose of this descriptive study was to describe the demographic characteristics of students who took the Georgia Alternate Assessment (GAA) and students who took the Georgia Milestones End of Grade Assessment (EOG) in English/Language Arts using publicly available data from the 2014-2015 academic year. Additionally, the study investigated disproportionality of certain student groups who took the GAA and EOG in English/Language Arts. A series of chi-square analyses resulted in significant overrepresentation of male, Black or African American, and economically disadvantaged students on the GAA. Our results indicated underrepresentation of migrant, limited English proficient, and Hispanic students on the GAA. Results of the present study have important implications for the student referral process and inclusion criteria for alternate assessment.

Keywords

Georgia Alternate Assessment, standardized testing, demographics, bias

Introduction

Alternate assessments were nationally mandated by the Individuals with Disabilities Education Act (IDEA) amendments of 1997 as a mechanism of inclusion for students with significant cognitive disabilities (Towles-Reeves, Kearns, Kleinert, & Kleinert, 2009). Additionally, the 2001 reauthorization of the No Child Left Behind Act (NCLB) recognized the need for alternate assessments for individuals with significant cognitive disabilities. Between 1-2% of students in the United States qualify for alternate assessments to replace large-scale educational evaluations (Evans & Domaleski, 2018; Harding, 2016). This is not surprising given that the prevalence of children diagnosed with cognitive or intellectual disability is approximately 1.48% among boys and .90% among girls (Zablotsky, Black, & Blumberg, 2017). Additionally, the prevalence of children diagnosed with intellectual disability is similar across racial and ethnic groups including Hispanic ethnicity, non-Hispanic black children, and non-Hispanic other children (Zablotsky, Black, & Blumberg, 2017). However, among all students receiving special education services under IDEA, the percentages of students who receive services significantly differs by race/ethnicity. Specifically, the percentage of students who received services for an intellectual disability was highest for those who were Black (9 percent) and the percentages for students of the other races/ethnicities ranged from 5 to 7 percent (National Center for Educational Statistics, 2019). Given these prevalence rates, we would expect proportionality among demographic groups on alternate assessments as compared to the general student population or perhaps those identified for special education.

Individualized Educational Plan (IEP) teams typically determine eligibility for special services under IDEA as well as whether or not students qualify to participate in statewide assessments (Thurlow, Lazarus, Thompson, & Morse, 2005). There is a significant amount of variability in the student referral process and inclusion criteria for alternate assessment (Albus & Thurlow, 2012; Saven, Anderson, Nese, Farley, & Tindal, 2016; Thurlow, 2005; Thurlow et al., 2005). Such variability may contribute to potential biases and disproportionate representation among students participating in alternate assessment. Unfortunately, disproportionality regarding alternate assessment is a reasonable concern given the plethora of literature documenting disproportionality among culturally and linguistically diverse students in special education (Cartledge, Kea, Watson, & Oif, 2016; Counts, Katsiyannis, & Whitford, 2018; Ostendorf & Choi, 2017) as well under-representation of these groups in enrichment programs (Lakin, 2016; Vega & Moore, 2008; Worrell, Subotnik, Olszewski-Kubilius, & Dixson, 2019). There is currently no research available on the disproportionality of demographic groups on alternate assessments in the state of Georgia.

Therefore, the purposes of this descriptive study were to (1) describe the demographic characteristics of students who took the Georgia Alternate Assessment (GAA) and students who took the Georgia Milestones End of Grade Assessment (EOG) in English/Language Arts and (2) investigate potential disproportionality of certain student groups the tests. In other words, are all student demographic groups (i.e., gender, race, socio-economic status, language proficiency, migrant status) proportionally represented on the GAA and the EOG?

Literature Review

A survey of all 50 states conducted by Thurlow, Lazarus, Thompson, and Morse (2005) determined the policies used in establishing participant guidelines of statewide assessments for students with disabilities. An overwhelming majority of states (49) allowed Individualized Educational Plan (IEP) teams to determine whether or not students participate in statewide assessments (Thurlow et al., 2005). To a much lesser extent, decisions were made based on whether or not the student was taught in course or content areas covered by the assessment (27 states), and 20 states also required parental involvement in the assessment participation decision-making process (Thurlow, 2005). A second study conducted by Albus and Thurlow (2012) echoed these sentiments but also added that 49 states mandated that a student must present with a significant cognitive disability to participate in alternate assessments. However, since no federal definition for the term 'significant cognitive disability' exists, states receive some degree of leniency in the implementation of participation guidelines (Saven et al., 2016). Likewise, although many states have chosen to use alternate standards for students that take alternate assessments (Evans & Domaleski, 2018; Streagle & Scott, 2015), some states like Georgia do not (Roach, Elliot, & Webb, 2005).

Furthermore, there are a number of exclusionary criteria specified in state alternate assessment guidelines. For example, 40 states do not allow specific disability labels (e.g., multiple disabilities, traumatic brain injuries) or descriptive characteristics of student challenges to be used as participation guidelines (Albus & Thurlow, 2012). Thirty-one states do not allow language, social, cultural, or economic differences to influence their decision making, and 30 states do not use extended absences or low achievement levels to dictate whether or not a student should be tested using alternate assessments (Albus & Thurlow, 2012). Although it is possible that policies have changed since 2012, a more current review of these is not available in the literature.

The Georgia Department of Education (GADOE, 2017a) released a short list of inclusion criteria for testing students using alternate assessment (see Appendix A). The IEP team may recommend general state assessments with or without accommodations or an alternate assessment only after carefully considering the inclusion criteria. Within the decision framework, if yes is answered to each of the criterion questions (e.g., "Does the student require intensive, individualized instruction in a variety of instructional settings?"), a student should participate in alternate assessment. However, within this process, students with significant cognitive disabilities may still be vulnerable to biases that teachers and school staff may hold pertaining to certain race, gender, and socio-economic status, even when they are held unintentionally (McGinnis, 2017).

Critical race theory posits that race and the potential for racism and "othering" is ever present in social settings, including education and the testing context (Ladson-Billings, 1998). Critical race theory has been used to examine the inclusion of people of color in curriculum and testing standards (Vasquez Heilig, Brown, & Brown, 2012), as well as testing as deficit-focused (Valencia, 2010). Thus, it is plausible that social interactions between teachers and students prior and during the test referral process may affect eligibility decisions and subsequent differences in demographics of those taking the alternate assessment (Cho & Kinston, 2015). Again, disproportionality regarding alternate assessment is a reasonable concern given the literature documenting disproportionality in special education and enrichment programs (Cartledge, Kea, Watson, & Oif, 2016; Counts, Katsiyannis, & Whitford, 2018; Lakin, 2016; Ostendorf & Choi, 2017; Vega & Moore, 2008; Worrell, Subotnik, Olszewski-Kubilius, & Dixson, 2019).

The importance of referring student with disabilities to the most fair and equitable method of performance assessment necessitates the exploration of disproportionality in alternate assessments (Albus & Thurlow, 2012; Harding, 2016; Harris, 2004; Salter, 2016). Therefore, the purpose of this descriptive study was to describe the demographic characteristics of students who qualified to take the Georgia Alternate Assessment (GAA) in English/Language Arts and investigate disproportionality of certain student groups. In other words, are all student demographic groups (i.e., gender, race, socio-economic status, language proficiency, migrant status) proportionally represented across test types, the GAA versus the Georgia Milestones End of Grade Assessment (EOG)? Although the demographics will not tell the stories of the individual minority students involved, they offer an exploratory snapshot into current practices that may prompt further investigation in the form of educational research and reform.

Methodology

Data Source

The data analyzed for this descriptive quantitative study is from the GAA and EOG archival data from the Georgia Governor's Office of Student Achievement (GOSA). The data for this study came from the 2014-2015 archives, which is the most recent data collection available for public use. This data collection is typically utilized by education stakeholders to examine education statistics for individual schools and districts across Georgia. The files accessed from GOSA included aggregate data of student demographics (i.e., gender, race/ethnicity, and special populations) and test type taken (i.e., GAA or EOG). Special population subgroups included migrant, limited English proficient, economically disadvantaged, and students with disabilities. The data were entered in SPSS 24 (IBM Corp., 2016), a statistical software, for analysis.

Participants

The 2014-2015 set of data from GOSA included a total of 11,012 students who were tested in English Language Arts using the GAA, and a total of 772,114 students who were tested in English Language Arts using the EOG. The data included students from a variety of counties across the state of Georgia. Table 1 shows the distribution of demographics based on the two test types.

Table 1
Number of Students Tested on the GAA and EOG by Demographic Summary

	Test Type	
Demographic	GAA (%)	EOG (%)
Gender		
Male	7,285 (66.2%)	393,368 (50.9%)
Female	3,727 (33.8%)	378,746 (49.1%)
Race/Ethnicity		
Black or African American	5,070 (47.4%)	282,866 (38.1%)
White	4,072 (38.1%)	321,504 (43.4%)
Hispanic	1,258 (11.8%)	110,563 (14.9%)
Two or More Races	285 (2.7%)	25,692 (3.3%)
American Indian or Alaskan Native	16 (0.2%)	1,429 (0.2%)
Native Hawaiian or Other Pacific Islander	0 (0%)	865 (0.1%)
Special Populations		
Migrant	15 (1.4%)	1,912 (0.2%)
Limited English Proficient	444 (4.0%)	58,036 (7.4%)
Economically Disadvantaged	8,341 (75.7%)	508,025 (65.8%)
Students with Disabilities	11,012 (100%)	91,350 (11.7%)
Total	11,012 (100%)	772,114 (100%)

Measures

Georgia Milestones End of Grade Assessment (EOG). Students in grades 3-8 take EOG assessments which consist of both selected- and constructed-response items across content areas (e.g, English Language Arts, Mathematics, Science, and Social Studies). The EOG has demonstrated acceptable reliability (ranging from alphas of .85-.93) and validity in measuring student achievement (see GADOE, 2016). This current study analyzed the most current data available, from the 2014-2015 school year, in which the state assessment measured academic growth toward the previously implemented Common Core Georgia Performance Standards (CCGPS). The CCGPS were adopted July 2010 and provided a set of consistent expectations across Georgia

and the nation (GADOE, 2010). However, despite setting ambitious goals for teaching and learning, the CCGPS fell out of favor with educators, parents, and community members (Gorman, 2016) and were replaced with the Georgia Standards of Excellence (GSE). During the 2015-2016 school term, GSE were implemented for English Language Arts and Mathematics; however, Science and Social Studies did not adopt GSE until the 2017-2018 school year. Furthermore, GSE in Fine Arts are required to be fully implemented during 2018-2019.

Georgia Alternate Assessment (GAA). The GAA is a portfolio-based evaluation that assess different areas of content for different grades. For example, Kindergarten and Grades 3, 4, 6, and 7 assemble a portfolio in English Language Arts and Mathematics. Grades 5, 8 and 11 assemble a portfolio in English Language Arts, Mathematics, Science, and Social Studies (GADOE, 2017b). The GAA is scored in four discrete dimensions: fidelity to standard, context, achievement/progress, and generalization (GADOE, 2013). Fidelity to standard includes the degree to which the student's work addresses the grade-level standard it is aligned to, and context is defined by the degree to which the student work exhibits the use of grade-appropriate materials in a purposeful and natural/real-world application (GADOE, 2013). Achievement/progress assesses the increase in the student's proficiency of skill across two collection periods, and generalization assesses the student's opportunity to apply the learned skill in other settings and/or with various individuals in addition to the teacher or paraprofessional (GADOE, 2013).

Debating reliability and validity of the GAA for measuring student achievement is beyond the scope of this study. However, there are some recognized limitations inherent in the portfolio-based measure. Teachers rate their students in each of the four dimensions using a rubric ranging from 1-4 points except for fidelity to standard which ranges from 1-3 points (Barge, 2013). Scoring inconsistencies pose a reliability risk as those assessing portfolio-based work may have differing expectations and interpretations of student work. The design of scoring rubrics and procedures for a relatively small and heterogeneous population across states that differ in educational standards is one of the greatest challenges in the development of alternate assessments (Elliot & Roach, 2007; Marion & Pellegrino, 2006). Additionally, some do not consider the GAA to be a valid assessment for students with significant cognitive disabilities because the GAA focuses on selected skills aligned to statewide Georgia curriculum standards instead of the necessary life skills needed outside the classroom (GADOE, 2017b; Harris, 2014; Salter, 2016). Teachers indicated that the GAA should be individualized to each student (Salter, 2016). However, while individualization would accurately measure the strengths and weaknesses of each student, it would be extremely difficult to

develop scoring methods for each student presenting with a significant cognitive disability.

Analysis

To analyze this data, a series of 10 Pearson's chi-square analyses were conducted. A Pearson's chi-square test is used to determine whether or not there is a relationship between any two categorical variables (Field, 2013). In this study, Pearson's chi-square was used to determine whether demographic data (i.e., gender, race/ethnicity, special population status) relates to the type of test that was administered (i.e., GAA or EOG). Cramer's V was used to determine the size or magnitude of the effect. Cramer's V is the most common strength test used to test data when a significant chi-square result has been obtained (McHugh, 2013). A weak effect of Cramer's V ranges from less than plus or minus 0.10 or from plus or minus 0.10 to 0.19. A moderate effect of Cramer's V ranges from plus or minus 0.20 to 0.29. A strong effect of Cramer's V ranges from more than plus or minus 0.30 (Jones, 2008).

To begin analyzing the data, subgroup demographics and test format variables were dummy coded. Next, each subgroup was formatted and coded as binary 0s and 1s. For example, if the GAA was administered, it received a dummy code of 1. The EOG received a dummy code of 0. One of the subgroup demographic variables that was examined also included migrant (1) and non-migrant (0) students. For each chi-square calculated, a separate SPSS data file was created with relevant subgroup information coded as binary 0s and 1s.

Once demographic variables and test format were dummy coded, frequency or the number of students tested with either the GAA or the EOG based on the specific demographic category were input. Descriptive statistics were then run on each of the demographic variables listed in the above paragraph through the use of crosstabs that generate information about bivariate relationships.

Results

Test Format and Demographic Variables

Figure 1 compares the percentage of students tested on the GAA to the percentage of students tested on the EOG within specific demographics groups. In terms of gender and test format, there was a significant association, χ^2 (1) = 1,005.00, $p < .001$. Based on the odds ratio (1.9), male students were 90% more likely to be tested on the GAA than female students. The effect size was weak, $V = .04$. There was a significant association between test format and economically disadvantaged students, χ^2 (1) = 478.38, $p < .001$. Based on the odds ratio (1.6),

economically disadvantaged students were 60% more likely to be tested on the GAA than students who were not economically disadvantaged. The effect size was weak, $V = .03$. In terms of English language proficiency and test format, there was a significant association, $\chi^2 (1) = 190.78$, $p < .001$. Based on the odds ratio (0.53), students with limited English proficiency were 47% less likely to be tested on the GAA than other students. The effect size was weak, $V = .02$. Similarly, there was a significant association between test format and migrant student status, $\chi^2 (1) = 5.49$, $p = .019$. Based on the odds ratio (0.57), migrant students were 43% less likely to be tested on the GAA than other students. The effect size was weak, $V = .003$.

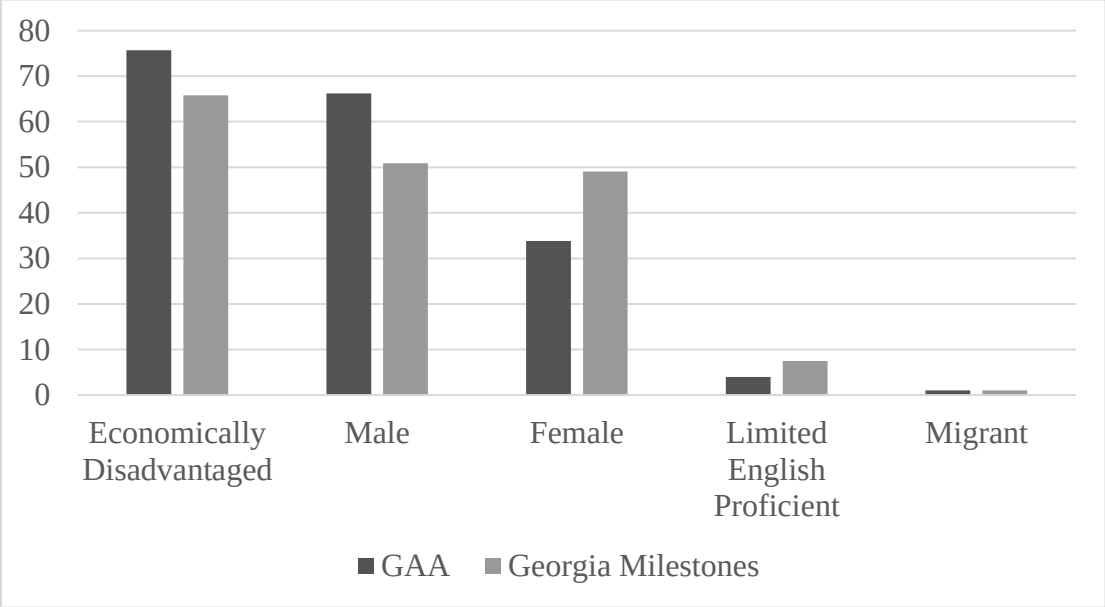

Figure 1. Percentage of students tested on the GAA and the Georgia Milestones within specific demographics group.

Note: Small percentages of migrant students were tested on the GAA (0.1%) and the Georgia Milestones (0.2%) which is why graph visibility is low. Results do indicate, however, that there is a significant association between test format and migrant student status, $\chi^2 (1) = 5.49$, $p = .019$, in which migrant students are 43% less likely to be tested on the GAA than the Georgia Milestones based on the odds ratio (0.57).

Test Format and Race/Ethnicity

Figure 2 compares the percentage of students tested on the GAA to the percentage of students tested on the EOG by race/ethnicity. There was a significant association between test format and Hispanic ethnicity, χ^2 (1) = 74.37, $p < .001$. Based on the odds ratio (0.73), Hispanic students were 27% less likely to be tested on the GAA than non-Hispanic students. The effect size was weak, $V = .01$. There was a significant association between test format and race, χ^2 (1) = 358.140, $p < .001$. Based on the odds ratio (1.3), White students were 30% more likely to be tested on the GAA than Asian students. Based on the odds ratio (1.18), White students were also 18% more likely to be tested on the GAA than students of two or more races. However, based on the odds ratio (1.38), Black or African American students were 38% more likely to be tested on the GAA than White students. Based on the odds ratio (1.8), Black or African American students were also 80% more likely to be tested on the GAA than Asian students. Based on the odds ratio (1.64), Black or African American students were also 64% more likely to be tested on the GAA than students of two or more races. However, based on the odds ratio (1.1), students of two or more races were 10% more likely to be tested on the GAA than Asian students. There was no significant association between test format and American Indians or Alaskan Natives student racial identity. Overall, the effect size of the association between test format and race was weak, $V = .02$.

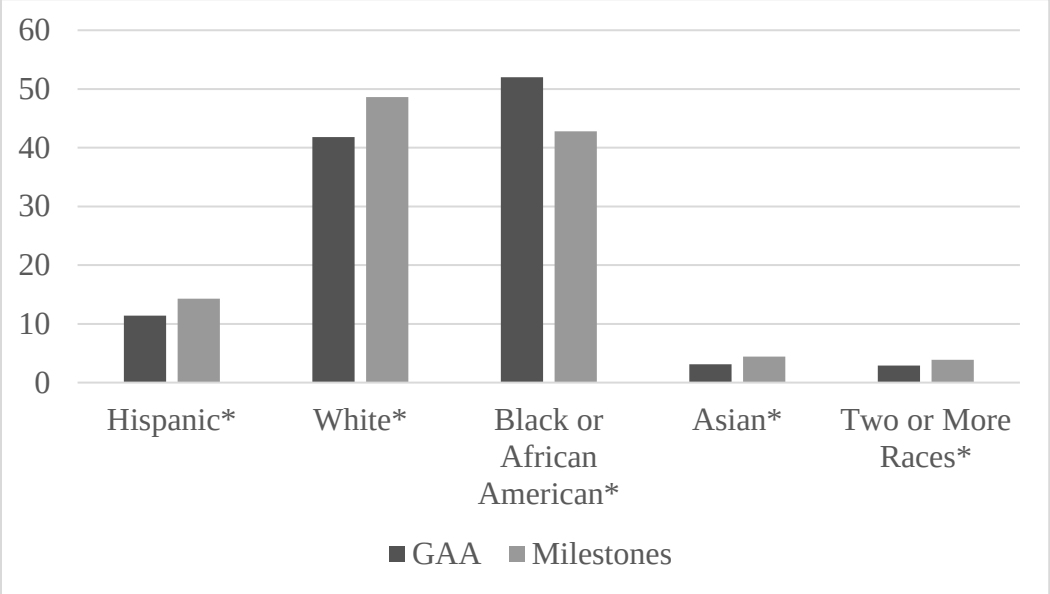

Figure 2. Percentage of students tested on the GAA and the Georgia Milestones within race/ethnicity.

Discussion

Between 1-2% of students with significant cognitive disabilities in the United States are assessed using alternate assessments such as the GAA, which is consistent with overall prevalence rates of cognitive or intellectual disability in the general student population (Evans & Domaleski, 2018; Harding, 2016; Zablotsky, Black, & Blumberg, 2017). But while the prevalence of children diagnosed with intellectual disability is similar across racial and ethnic groups (Zablotsky, Black, & Blumberg, 2017), we do not observe proportionate trends among all students receiving special education services under IDEA, specifically those with intellectual disability (National Center for Educational Statistics, 2019) leading to reasonable concerns about disproportionality among demographic groups on alternate assessments. There is currently no other research available on the disproportionality of demographic groups on alternate assessments in the state of Georgia. Therefore, the purposes of this descriptive study were to (1) describe the demographic characteristics of students who took the Georgia Alternate Assessment (GAA) and students who took the Georgia Milestones End of Grade Assessment (EOG) for English/Language Arts and (2) investigate potential disproportionality of certain student groups the tests.

Given what we know about prevalence rates, we might expect proportionality among demographic groups on alternate assessments compared to diagnostic rates of students with significant cognitive disability in the general student population or perhaps among those identified for special education. This descriptive study serves as one of the first steps of inquiry into a new topic. Although the results of the present study are weak in magnitude, the findings are significant, suggesting disproportionality across demographic categories and test format in the state of Georgia. Furthermore, cognitive explanations within critical race theory, including in-group favoritism and referral bias, may offer tentative explanations for the group differences observed.

Gender

In this study, male students were 90% more likely to be tested on the GAA than female students. Conversely, females were overrepresented on the EOG. Findings are consistent with prevalence rates indicating boys are identified with intellectual disability more often than girls (Zablotsky, Black, & Blumberg, 2017). Findings are also consistent with research conducted by Anderson (1997) who found a large discrepancy between the numbers of females to males in learning disabled programs. Still, a 90% greater likelihood seems higher than expected. Referral bias may offer one possible explanation for the results of this study. Referrals for special education determine which students are eligible to receive special education services including accommodations in instruction,

accommodations in testing, or evaluation through the use of alternate assessments (Anderson, 1997). Gender referral bias may be associated with student disruption in the classroom and stereotypical behaviors based on gender (Anderson, 1997). Most teachers find that boys are more active, more inattentive, and have more problems in behavior, language, and academics than their female peers (Anderson, 1997). Therefore, boys are more likely to let their frustration show in a classroom setting which disrupts learning and instruction and provides teachers with an incentive to solve the issue by referring them to special education services or in this case, alternate assessment. It seems plausible that biased perceptions based on gender may also influence referral for alternate assessment.

Race and Socio-economic Status

Results also indicated a large overrepresentation of Black or African American groups and economically disadvantaged students taking the GAA. Conversely, these minority groups were underrepresented on the EOG. Black or African American students were 38% more likely to be tested on the GAA than White students, 80% more likely to be tested on the GAA than Asian students, and 64% more likely to be tested on the GAA than students of two or more races. This is consistent with special education statistics indicating the percentage of students who received services for an intellectual disability was highest for those who were Black as compared to other racial and ethnic groups (National Center for Educational Statistics, 2019). Similarly, economically disadvantaged students were 60% more likely to be tested on the GAA than other economically advantaged students. Research indicates that socioeconomic status is a confounder for disparities among mild intellectual disabilities but appears to be less influential on more serious intellectual disabilities (National Academies of Sciences, Engineering, and Medicine, 2015). Since group-based inequalities are still evident in education, in-group favoritism might account for the differences that we observed (Dasgupta, 2004; Glock & Böhmer, 2018). In-group favoritism is the tendency to evaluate one's own membership group (the in-group) or its members more favorably than a non-membership group (the out-group) or its members (Hewstone, .Rubin, & Willis, 2002). Intergroup attitudes may include unconscious, unintentional, or implicit bias where well-intentioned people are largely unaware of their favoritism (Hewstone, Rubin, & Willis, 2002).

Although the racial diversity of teachers, administrators, and other IEP team members is on the increase, a majority (82%) of public-school teachers are still White and non-migrant (Deruy, 2018). Dasgupta (2004) found that White Americans show strong implicit preference for their own group and relative bias against African Americans, Latinos, Asians, and other non-Americans. Teachers by profession also fall within a low to middle class socio-economic status.

Therefore, demographic differences in test takers of the EOG and the GAA may be the result of unintentional in-group favoritism. When members of an in-group prefer their in-group over others, a separation occurs that divides already disadvantaged populations even further. Therefore, less resources are made available and less support is given to cultivate a child's education. Future research should focus on the overrepresentation of Blacks or African Americans and economically disadvantaged students on the GAA. A more comprehensive study of teacher demographics should be observed to determine whether in-group favoritism and referral bias are factors in these differences.

Primary Language/Culture

Conversely, our results indicated that migrant, limited English proficient, and Hispanic students are underrepresented in the GAA. Even in cases of underrepresentation, implicit associations held by teachers and staff can contribute to racial inequality via racial bias (Warikoo, Sinclair, Fei, & Jacoby-Senghor, 2016). Perhaps educators are less equipped to detect cognitive disabilities in migrant, limited English proficient, and Hispanic populations due to language barriers or other cultural differences. More research is needed to determine why there is an underrepresentation of the above populations in the GAA. Some key questions that future research needs to address include: Can language barriers or other cultural differences explain why students are referred less often to take alternate assessments? Are IEP team members less equipped to identify cognitive disabilities in these populations due to cultural factors or language barriers? Another avenue for future research could explore how our student support teams are being trained and prepared to initially identify and meet the needs of intellectually disabled populations.

Limitations

This study examined test format and demographics categories in the state of Georgia and as with any study, limitations should be noted. First of all, only the English/Language Arts content area was analyzed. Trends could differ across content areas. Although some of the discrepancies across groups may be due to actual differences in performance and/or ability rather than bias in the referral or assessment, the differences are significant and should be further examined. An important limitation of the present study is that since all of our variables are nominal, we cannot determine a cause and effect relationship. Therefore, our findings and possible explanations are open to interpretation instead of determining causation. Additionally, chi-square studies, such as this one, are also highly sensitive to sample size. As sample size increases, absolute differences become a smaller proportion of the expected value. With large sample sizes on

both the GAA (11,012) and the EOG (91,350), our findings of demographics differences in test format may in part be due to large sample size.

Summary and Implications

The present study described student demographic groups using the most recent GAA data available to the public and determined that some groups of students are referred to alternate testing more than others. Demographic group members were disproportionately represented across test types administered (i.e., EOG and GAA). Males, African Americans, and economically disadvantaged students were more likely to be assessed with the GAA for English/Language Arts. Additionally, migrant, limited English proficient, and Hispanic students were less likely to be assessed with the GAA. In conclusion, this study found evidence of disproportionality among the students who took the alternate assessment; however, the explanation for these findings are yet to be understood (Hutchison, 2018).

Results of the present study have important implications for the testing of all students regardless of disability characteristic or demographics category. If referral bias or in-group favoritism explain our results, more educator training to combat these discrepancies should be enacted (e.g., Banakou, Hanumanthu, & Slater, 2016). The main purpose of these trainings would be to educate teachers, other IEP team members, and parents on implicit biases and how such biases influence educational decisions. Additional procedural safeguards should be established within the referral process to minimize overrepresentation or underrepresentation of identified groups (Cho & Kingston, 2015). Although not within the scope of this study, reliability and validity of the portfolio based GAA for students with significant cognitive disabilities is of concern and future researchers may wish to better address these measurement issues (Saven et al., 2016).

References

Albus, D. & Thurlow, M. (2012). Alternate assessments based on alternate achievement standards (AA-AAAS) participation policies (Synthesis Report 88). *Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes*. Retrieved from https://nceo.umn.edu/docs/OnlinePubs/Synthesis88/SynthesisReport88.pdf

Anderson, K. (1997). Gender bias and special education referrals. *Annals of Dyslexia*, *47*(1), 151-162. doi: 10.1007/s11881-997-0024-8

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in Human Neuroscience*, *10*, 601.

Barge, J. (2013). Score interpretation guide. Retrieved from https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/GAA%20Score%20Interpretation%20Guide%202013.pdf

Cartledge, G., Kea, C. D., Watson, M., & Oif, A. (2016). Special education disproportionality: A review of response to intervention and culturally relevant pedagogy. *Multiple Voices for Ethnically Diverse Exceptional Learners*, *16*(1), 29-49.

Cho, H. J., & Kingston, N. (2015). Examining teachers' decisions on test-type assignment for statewide assessments. *The Journal of Special Education*, *49*(1), 16-27.

Counts, J., Katsiyannis, A., & Whitford, D. K. (2018). Culturally and linguistically diverse learners in special education: English learners. *NASSP Bulletin*, 102(1), 5-21.

Dasgupta, N. (2004). Implicit ingroup favoritism, outgroup favoritism, and their behavioral manifestations. *Social Justice Research*, *17*(2), 143-169. doi: 10.1023/B:SORE.0000027407.70241.15

Deruy, E. (2018). Student diversity is up but teachers are mostly White. Retrieved from https://aacte.org/news-room/aacte-in-the-news/347-student-diversity-is-up-but-teachers-are-mostly-White

Elliot, S. & Roach, A. (2007). Alternate assessments of students with significant disabilities: Alternative approaches, common technical challenges. *Applied Measurement in Education*, *20*(3), 301-333. doi: 10.1080/08957340701431385

Evans, C. M., & Domaleski, C. (2018). Guidance for examining district alternate assessment participation rates. *National Center on Educational Outcomes*. Retrieved from https://eric.ed.gov/?id=ED591939

Georgia Department of Education. (2010). *Common core Georgia performance standards (CCGPS)*. Retrieved from

http://archives.doe.k12.ga.us/DMGetDocument.aspx/CCGPSJuly8.pdf?p=6CC6799F8C1371F6538E425D6722145B3F25DDC7CA674A057392A682351E2925

Georgia Department of Education. (2013). *Georgia alternate assessments (GAA): Questions and answers for parents of Georgia students*. Retrieved from http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/2012-2013%20Informational%20Brochure.pdf

Georgia Department of Education. (2015). *Georgia standards of excellence (GSE).* Retrieved from https://www.georgiastandards.org/Georgia-Standards/Pages/default.aspx

Georgia Department of Education. (2016). *Validity and reliability for the 2014-2015 Georgia milestones assessment system.* Retrieved from https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Milestones/2014-15_Georgia_Milestones_Validity_and_Reliability_Brief.pdf

Georgia Department of Education. (2017a). *Guidelines for participation in the GAA.* Retrieved from http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/GAA/GAA_16-17/GAA_Participation_Graphic.pdf

Georgia Department of Education. (2017b). *The Georgia alternate assessment (GAA).* Retrieved from http://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/GAA.aspx

Gibbons, J. D. & Chakraborti, S. (2003). *Nonparametric statistical inference.* 4th ed. New York: Marcel Dekker.

Glock, S., & Böhmer, I. (2018). Teachers' and preservice teachers' stereotypes, attitudes, and spontaneous judgments of male ethnic minority students. *Studies in Educational Evaluation*, *59*, 244-255.

Gorman, N. (2016). How the common core became a misbranded buzzword used to define federal overreach. *Education World*. Retrieved from http://www.educationworld.com/a_news/how-common-core-became-misbranded-buzz-word-used-define-federal-overreach-207237223

Harding, I. L. (2016). *Alternate assessment of students with significant cognitive disabilities: A research report*. Xlibris Corporation.

Harris, V. (2014). Special education teachers' experiences and perceptions of the GAA. *Electronic Theses & Dissertations*. 1062. Retrieved from https://digitalcommons.georgiasouthern.edu/etd/1062

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*, 575-604. doi: 10.1146/annurev.psych.53.100901.135109

Hutchison, C. B. (2018). Re-thinking disproportionality in special education as a self-fulfilling prophecy. *Insights into Learning Disabilities*, *15*(2), 113-116.

IBM Corp. (2016). *IBM SPSS Statistics for Windows*, Version 24.0. Armonk, NY: IBM Corp.

Jones, L. (2008). *Tests of association for nominal data*. Retrieved from https://www.angelo.edu/faculty/ljones/gov3301/block14/objective3.htm

Ladson-Billings, G. (1998). Just what is critical race theory and what's it doing in a nice field like education?. *International Journal of Qualitative Studies in Education*, *11*(1), 7-24.

Lakin, J. M. (2016). Universal screening and the representation of historically underrepresented minority students in gifted education: Minding the gaps in Card and Giuliano's research. *Journal of Advanced Academics*, *27*(2), 139-149.

Marion, S. & Pellegrino, J. (2006). A validity framework for evaluating the technical quality of alternate assessments. *Educational Measurement: Issues and Practice*, *25*(4), 47-57. doi: 10.1111/j.1745-3992.2006.00078.x

McGinnis, C. (2017). Effects of implicit bias on teacher's expectations of student relationships. *Public Access Theses and Dissertations from the College of Education and Human Services*. 294. Retrieved from http://digitalcommons.unl.edu/cehdiss/294

McHugh, M. (2013). The chi-square test of independence. *Biochemia Medica*, *23*(2), 143-149. doi: 10.11613/BM.2013.018

National Center for Educational Statistics (2019). *Status and trends in the education of racial and ethnic groups*. Retrieved from https://nces.ed.gov/programs/raceindicators/indicator_RBD.asp.

National Academies of Sciences, Engineering, and Medicine (2015). *Mental disorders and disabilities among low-income children*. Washington, DC: The National Academies Press. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK332882/pdf/Bookshelf_NBK332882.pdf.

Ostendorf, R. J., & Choi, N. (2017). Cultural and linguistic diversity in special education in the United States and South Korea: Exploring current practices and recommendations. *International Journal of Special Education*, *32*(4), 793-808.

Roach, A. T., Elliott, S. N., & Webb, N. L. (2005). Alignment of an alternate assessment with state academic standards: Evidence for the content validity of the Wisconsin alternate assessment. *The Journal of Special Education*, *38*(4), 218-231.

Salter, M. (2016). Secondary special education teachers' perceptions of the GAA (GAA). *Tift College of Education, Mercer University.* Retrieved from

https://search.proquest.com/openview/e1f9d7a79202637adfaad30e4fcda1ec/1?pq-origsite=gscholar&cbl=18750&diss=y

Saven, J. L., Anderson, D., Nese, J. F., Farley, D., & Tindal, G. (2016). Patterns of statewide test participation for students with significant cognitive disabilities. *The Journal of Special Education*, *49*(4), 209-220.

Solorzano, D. G., & Yosso, T. J. (2001). From racial stereotyping and deficit discourse toward a critical race theory in teacher education. *Multicultural Education*, *9*(1), 2.

Streagle, K., & Scott, K. W. (2015). The alternate assessment based on alternate achievement standards eligibility decision-making process. *The Qualitative Report*, *20*(8), 1290-1312.

Thurlow, M. L., Lazarus, S. S., Thompson, S. J., & Morse, A. B. (2005). State policies on assessment participation and accommodations for students with disabilities. *The Journal of Special Education*, *38*, 232–240. doi: 10.1177/00224669050380040401

Towles-Reeves, E., Kearns, J., Kleinert, H., & Kleinert, J. (2009). An analysis of the learning characteristics of students taking alternate assessments based on alternate achievement standards. *The Journal of Special Education*, *42*(4), 241-254. doi: 10.1177/0022466907313451

Valencia, R. R. (2010). *Dismantling contemporary deficit thinking: Educational thought and practice*. New York: Routledge.

Vasquez Heilig, J., Brown, K., & Brown, A. (2012). The illusion of inclusion: A critical race theory textual analysis of race and standards. *Harvard Educational Review*, *82*(3), 403-424.

Vega, D., & Moore III, J. L. (2018). Access to gifted education among African-American and Latino males. *Journal for Multicultural Education*, *12*(3), 237-248.

Warikoo, N., Sinclair, S., Fei, J., & Jacoby-Senghor, D. (2016). Examining racial bias in education: A new approach. *Educational Researcher*, *45*(9), 508-514.

Worrell, F. C., Subotnik, R. F., Olszewski-Kubilius, P., & Dixson, D. D. (2019). Gifted students. *Annual review of psychology*, *70*, 551-576.

Zablotsky, B., Black, L.I., & Blumberg, S.J. (2017). *Estimated prevalence of children with diagnosed developmental disabilities in the United States, 2014–2016*. NCHS Data Brief, no 291. Hyattsville, MD: National Center for Health Statistics.

Appendix. GAA Participation Guidelines

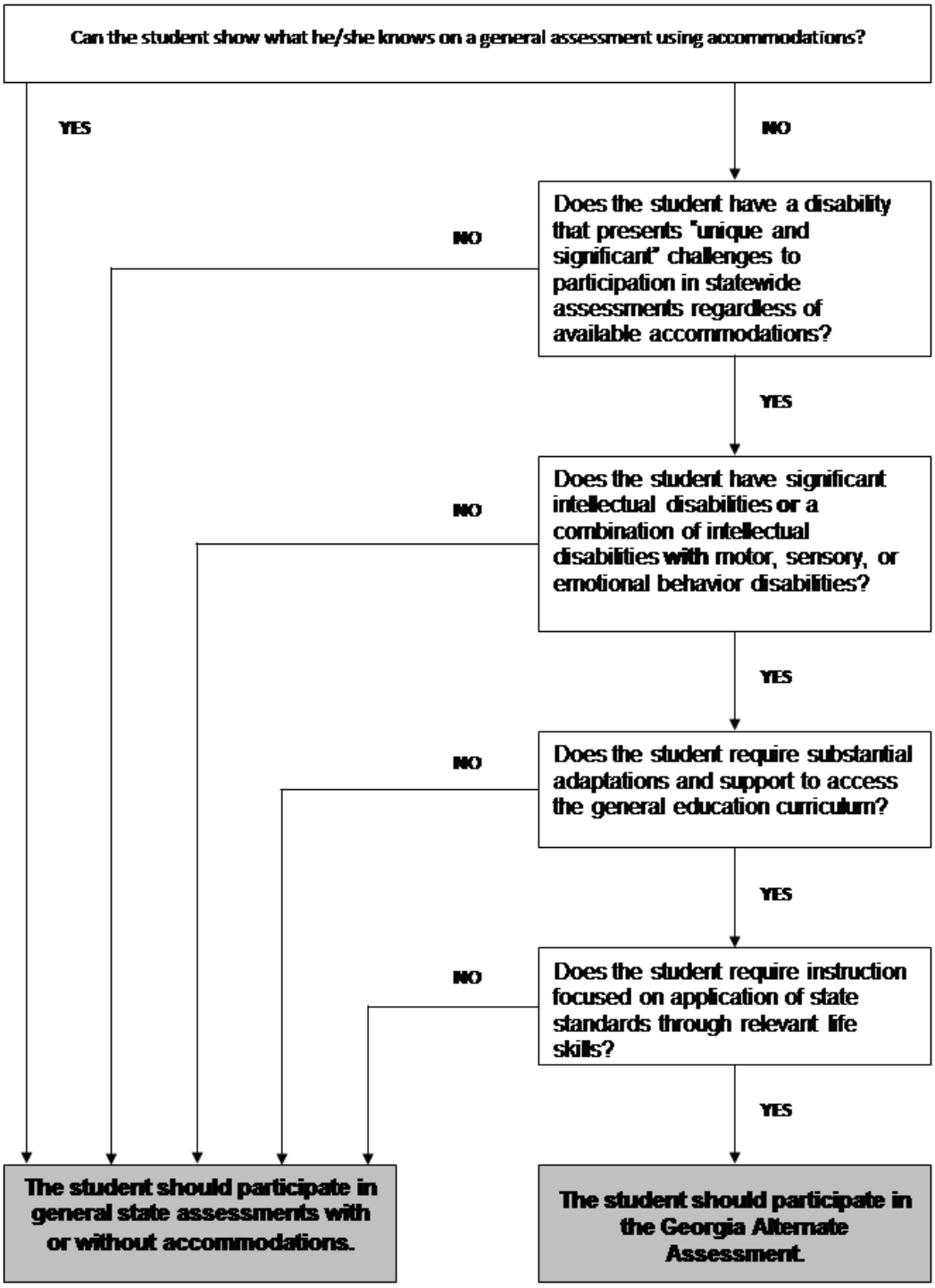

Figure A1. Guidelines for Participation in the GAA (GADOE, 2017).