

FROM FIXED TO FLEXIBLE: NEEDED CONDITIONS TO PROMOTE ELEMENTARY TEACHERS' EQUITABLE USE OF WITHIN-CLASS ABILITY GROUPING

SARAH CAROLEO

PH.D. STUDENT

SCHOOL OF EDUCATION, JOHNS HOPKINS UNIVERSITY

ORCID 0000-0002-0041-9616

ABSTRACT

Within-class elementary grouping is a staple of modern elementary instruction, as it ideally provides a structure in which classroom teachers can better manage academic diversity. However, it is often implemented ineffectively and/or inequitably due to various structural, cultural, and political features of school systems and teacher training programs. In this essay, I seek to delineate solutions via flexible grouping that combat historical inequities associated with student ability grouping, ultimately to equip teachers to both manage academic diversity and ensure that all students receive appropriately challenging instruction each day. The arguments put forth are informed by my ten years of work as an educator, instructional coach, and researcher, in which I have witnessed a strong, practical need for elementary small group instruction but have also grappled with how ability grouping often inequitably sorts and fixes students into groups that fuel de facto tracking. This work will benefit school and district leaders and teacher preparation programs, as they seek to address systemic issues related to teachers' ability grouping practices. Most importantly, it will provide tangible strategies and descriptions that equip elementary educators to leverage more flexible, equitable grouping practices in their classrooms.

Key Words: Ability grouping; elementary teaching; differentiation; flexible grouping; small group instruction.

Author Note

*I would like to extend my sincere thanks to Dr. Julia Burdick-Will, who provided initial guidance on the structure and framing of this piece, and Dr. Jonathan Plucker, who gave helpful feedback for revisions and edits. I would also like to express my gratitude to the editorial board of *The Neag School of Education Journal*, who supplied constructive and supportive feedback that expanded and strengthened the essay.*

INTRODUCTION

Elementary teachers are typically tasked with teaching all subjects to a body of students displaying highly varied levels of readiness for grade-level content, prompting many to utilize within-class small group instruction to level core reading and math instruction (Esposito, 1973; Sørensen, 1970). Research outlines the potential benefits of within-class ability grouping; it helps teachers provide differentiated instruction to academically diverse students (Adelson & Carpenter, 2011; Adodo & Agbayewa, 2011; Castle et al., 2005; Slavin, 1987) and become more familiarized with students' unique personalities and learning dispositions (Altintas & Ozdemir, 2015). Within-class ability grouping can also increase students' engagement since they interact more with the teacher and peers in small groups (Altintas & Ozdemir, 2015; Castle et al., 2005), as well as increase students' self-regulated behaviors (e.g., monitoring personal progress and talking about thinking) since learning is often more active in small groups (Stright & Supplee, 2002). In his best-evidence synthesis, Slavin (1987) found that within-class ability group instruction can have positive effects on student achievement when three criteria are met: (1) the grouping is based upon targeted skill differences of focus across students; (2) teachers flexibly move students based on current levels of understanding; and (3) teachers alter the pace and level of group-level instruction to correspond to students' readiness and rate of learning. A more recent second-order meta-analysis (Steenbergen-Hu et al., 2016) drew promising findings, as well; across 13 meta-analyses about ability grouping, within-class ability grouping had positive and significant effects (with effect sizes ranging from 0.19 to 0.30) on all students' subsequent academic achievement, regardless of initial ability level.

A notable concern often arises in this deeply normed model, however; teachers frequently do not receive sufficient—if any—pre-service or in-service training and resources that equip them to meet any of Slavin's (1987) three criteria (Fitzgerald et al., 2021; Harris, 2010; Tomlinson et al., 1994). To effectively and equitably group students for small group instruction, teachers need to acquire pedagogical/psychological knowledge and hold a range of skills that extend from that knowledge (Voss et al., 2011): how to evaluate student strengths, needs, and misconceptions related to a unit of study; how to match optimal peer groups for each topic; how to meaningfully differentiate instruction around the content standard; how to know when to move a student to a different group; how to appropriately challenge all learners in every group; and more. Structural, cultural, and political dynamics of schools often leave teachers undertrained in this knowledge and these skills, under-resourced to implement differentiated instruction, and pressured to group students in fixed ways that contribute to de facto segregation patterns (Buttaro et al., 2010).

For example, without needed training and resources, many elementary teachers utilizing within-class ability grouping tend to disproportionately assign students from lower social classes, students with perceived behavioral challenges, students with disabilities, and students of color into lower ability groups and provide decontextualized instruction at a slower pace with conveyed low expectations in those groups (Becton, 2018; Calarco, 2014; Eder, 1981; Esposito, 1973; Oakes, 1992). These inequitable grouping patterns fuel early learning gaps and poor academic self-concept that negatively affect later achievement of students fixed into "low" ability groups from a young age (Plucker & Peters, 2016; Tyson, 2011).

In this essay, I provide historical context for how ability grouping has evolved in American K-12 classrooms, articulate why flexible grouping practices are more effective and equitable than fixed grouping practices, and synthesize how teacher training programs and school leaders can make structural changes to better support teachers' facilitation of flexible grouping. Until intentional, systemic action is taken to move elementary teachers away from the use of fixed ability grouping, American schools will continue to foster opportunity and performance gaps that harm historically marginalized groups of students. Thus, it is imperative to support and guide elementary educators in the successful use and implementation of flexible ability grouping.

THE HISTORY OF STUDENT GROUPING IN THE UNITED STATES

To avoid and repair the problematic aspects of within-class ability grouping in modern elementary settings, it is critical to first understand how grouping practices have manifested in American classrooms over time—and what they have meant for the outcomes of historically marginalized students. Prior to the 1820s, children of American colonizers largely learned from family and community members either at home or in one-room schoolhouses (Tyack, 1974). However, in the 1820s, a mass of immigrants entered the country, making it increasingly harder to serve a large range of children across all ages in one-room settings. Around the 1840s, formal public schools were ideated and formed across the country; by the 1860s, age-based grades were established in most schools to foster more homogeneity of developmental needs in classrooms. This structure proved helpful in accounting for the range of academic diversity among students. From 1900 to 1920, the United States experienced another boom in immigration, meaning that class and school sizes grew quickly in number again (Ellison & Hallinan, 2004; Worthy, 2010). Simultaneously, the eugenics movement, a classist, ableist, and White supremacist movement committed to elevating “genetically superior” individuals and families in society and separating “genetically inferior” from public spaces, was on the rise and bled into schooling policies and practices (Brookwood, 2021).

In the 1910s, school systems began employing homogenous grouping strategies beyond age-based grade levels in schools (Goldberg et al., 1966), likely motivated by both the practical need to address the sharp influx in student enrollment and the eugenicist philosophy that certain people, typically those who were Western European, needed to be equipped for certain roles in society—thus, grouping students according to “ability” was a natural means to prepare children for their “deserved” roles. In elementary schools, within-class ability groups emerged, in which teachers used small group instruction with relatively homogeneous groups of students, particularly for reading (Barr & Dreeben, 1983; Ireson & Hallam, 1999). In high schools in the 1930s, tracking, a form of between-class grouping, arose as a dominant grouping strategy (Ellison & Hallinan, 2004; Oakes, 1985). Tracking was a practice in which students were assigned to a certain level—vocational, general, or academic—based on their past school achievement and/or their Intelligence Quotient (IQ) score, which derived from tests developed by leading eugenicists who endorsed separating “superior” children from “inferior” peers in schools (Brookwood, 2021; Steenbergen-Hu et al., 2016). Students then primarily completed coursework associated with their distinct track and remained fixed in their tracks for the duration of their schooling experience (Oakes, 1985).

Initial research around these forms of ability grouping suggested that they increased student achievement. However, Goldberg et al. (1966) published an article that highlighted how most prior studies had not accounted for several confounding factors, such as class size, number of students involved, teaching methods, and more. Subsequently, researchers began accounting for these confounders in their statistical models and often found either null or negative effects of tracking and ability grouping for students placed within “average” and “low” ability groups (Barr & Dreeben, 1983; Eder, 1981; Esposito, 1973; Gamoran, 1986; Oakes, 1985), meaning that performance gaps widened even further as “high” achieving students were the only group who benefitted academically from tracking and within-class grouping. On average, students from low-income backgrounds, students of color, and students with disabilities disproportionately comprised these “low” ability groups (Esposito, 1973; Oakes, 1985)—fulfilling the eugenicist aim of separating students by class, race, and “ability.”

In elementary settings, several studies found that instruction for the designated “low” ability groups was often facilitated at a slower pace and focused more time on decontextualized skills (Allington, 1983; Eder, 1981; Gambrell et al., 1981). Teachers were more likely to hold and convey lower expectations and negative feelings toward students in “low” groups (Eder, 1981; Good & Brophy, 1972), showing more concern for managing their behavior than providing appropriately challenging instruction (Eder, 1981)—practices likely driven by implicit and/or explicit teacher bias against students from historically marginalized backgrounds. Students remained relatively fixed in their assigned ability groups; Barr and Dreeben (1983) found that 70% of first graders remained in the same reading group throughout the duration of their school year. In high schools, researchers found cumulative effects of students’ previous ability group assignments and their differential access to quality instruction on their achievement; studies suggested that their placements reinforced and exacerbated their initial perceived ability differences (Alexander & Cook, 1982; Ellison & Hallinan, 2004; Gamoran, 1989; Ireson & Hallam, 1999). Scholars also voiced concerns for the negative effects on qualitative aspects of students’ lives, such as motivation, academic self-concept, and confidence, that might arise when certain students are primarily fixed into lower groups or tracks (Esposito, 1973; Oakes, 1992; Oakes et al., 1997).

Although several studies indicated that students’ group placements were most strongly associated with their initial achievement scores, students’ socioeconomic status was often significantly associated with their group placements (Esposito, 1973; Steenbergen-Hu et al., 2016; Worthy, 2010), suggesting that this structure may have served as a capital-reproducing mechanism for students who entered the K-12 system with privilege (Bourdieu, 1973). Some noted how the organizational structure of tracking seemed to perpetuate de facto segregation of students both racially and socioeconomically (Buttaro et al., 2010; Eder, 1981; Oakes, 1985,1992), launching a movement that urged schools to detrack their organizational structures and point resources toward providing all children with high-quality instruction. By the mid-1990s, most schools had minimized or eliminated tracking and within-class grouping (Steenbergen-Hu et al., 2016), and research surrounding the effects of ability grouping faded (Worthy, 2010).

ABILITY GROUPING IN MODERN ELEMENTARY CLASSROOMS

At the turn of the 21st century, with a stronger push for accountability via the passage of the No Child Left Behind Act and increased standardized testing, the use of within-class ability grouping gained new saliency in elementary settings (Ireson & Hallam, 1999; Steenbergen-Hu, 2016; Tomlinson, 2000). In 1998, 29% of elementary students participated in ability groups as part of reading instruction; by 2009, that number jumped to 71% (Loveless, 2009). Likewise, by 2011, 61% of elementary students participated in ability groups for mathematics instruction (Steenbergen-Hu et al., 2016). It makes sense that this structure became useful in a testing-driven context, as homogeneous grouping can allow teachers to better adjust instruction to match students' current needs in relation to grade level standards and thus ameliorate external pressure for students performing just below grade level (i.e., "bubble kids") to meet grade level standards by testing time. Nationally representative data suggested that some overall progress was made in elementary grades during this time of increased accountability and within-class ability grouping, with the National Assessment of Educational Progress (NAEP) average 4th grade reading scale scores significantly moving from 213 to 221 (National Center for Education Statistics, 2012); however, those same growth patterns did not occur for students with diagnosed learning disabilities (Becton, 2018). The concept of differentiated instruction (DI) grew in popularity in both research and school settings with this new boom of within-class ability grouping (Altintas & Ozdemir, 2015; Pozas et al., 2020; Santangelo & Tomlinson, 2012; Tomlinson, 2000), bringing much-needed discussions about how to innovate instruction and better match instruction to students' current Zones of Proximal Development (Vygotsky, 1978).

Some equity issues persist with this modern form of within-class ability grouping, though. For example, using ECLS-K data, Buttaro et al. (2010) found that within-class ability grouping in kindergarten was more frequently used in schools with higher levels of racial and socioeconomic diversity and proportions of students of color; conversely, majority-White schools were the least likely to use within-class grouping. This finding suggests that this structure may still function as a mechanism that inequitably separates students within schools or distills instruction for students wrongly assumed to be incapable of handling grade-level content. Students from low-income backgrounds and students of color continue to be disproportionately assigned to "low" ability groups (Adodo & Agbayewa, 2011; Buttaro et al., 2010; Condrón, 2007), and students placed in "high" ability groups still enjoy better academic and social gains from participating in small group instruction than those placed in "low" groups (Bradbury, 2018; Buttaro et al., 2010; Castle et al., 2005; Marks, 2013; Plucker & Peters, 2016). Some research suggests that within-class ability grouping has yet to provide positive effects for students with disabilities (Becton, 2018), likely because they are over-placed and fixed into low ability groups, and teachers may hold or convey differential, biased expectations of them.

When considering both the pitfalls and benefits of elementary within-class grouping over time, a few themes emerge. First, fixed placements into ability groups or tracks typically lead to inequitable distribution of instructional time and resources, benefitting those who already hold privilege prior to entering kindergarten (Bourdieu, 1973; Oakes et al., 1997; Tomlinson, 2000); thus, any form of within-class grouping that employs fixed grouping practices should be eliminated as much as possible from modern elementary education settings.

Next, labelling practices—where educators label a student as “low”, “average”, or “high”, either implicitly or verbally to other adults—can reinforce notions of fixed ability, which may lead educators to make assumptions about where students belong and of how much they can achieve (Barr & Dreeben, 1983; Eder, 1981; Tyson, 2011). Even if unintentional, these assumptions can greatly impact students’ academic self-concepts and access to prerequisite content needed for secondary and postsecondary success (Buttaro et al., 2010; Corbett Burriss et al., 2008; Oakes, 1992; Tyson, 2011). Therefore, educators should be mindful of the labels they assign to children in relation to ability; question their own assumptions about ability, its origins, and its development; and reflect upon how those assumptions affect their teaching of their students.

Finally, when grouping is implemented in a flexible way that builds homogeneity around the target skills of focus, as Slavin (1987) originally found, within-class ability grouping can support teachers in accounting for wide ranges of academic diversity in their classroom—a challenge that often arises in elementary settings. However, the vision of all American teachers using this model flexibly has not yet been realized and requires intentional action to progress.

COMPARING FIXED AND FLEXIBLE GROUPING PRACTICES

As discussed above, within-class ability grouping as an instructional model pervades modern elementary classrooms (Steenbergen-Hu et al., 2016), and its use can vary significantly based on teachers’ training, teaching philosophies, management practices, and personal capacities (Adelson & Carpenter, 2011; De Neve et al., 2015; Chandra Handa, 2020). To increase the efficacy and equity of within-class ability grouping, teacher preparation programs, school instructional leaders, and teachers must become familiarized with the differences between fixed and flexible grouping and reflect on to what extent their classroom and school norms are fixed or flexible. Table 1 serves as a quick guide for comparison, and elaborations are provided below.

Table 1

Characteristics of Fixed Grouping versus Flexible Grouping

<u>Fixed Grouping</u>	<u>Flexible Grouping</u>
1. Groups are formed with the use of standardized or benchmark test results. (Fitzgerald et al., 2021)	1. Groups are formed with the use of frequent diagnostic and formative assessment data. (Borland et al., 2002; Castle et al., 2005; Tomlinson et al., 2003)
2. Students are considered and labelled high-, average-, or low-performing. (Fitzgerald et al., 2021)	2. Students’ current strengths and next steps for growth are routinely considered. (Tomlinson et al., 2003)
3. Little to no movement across groups occurs. (Missett et al., 2014; Tomlinson et al., 1997)	3. Students are regularly moved to a group with their shared goal for the day or week. (Tomlinson et al., 2003)

4. Student behaviors inform their group placement. (Kim et al., 2020; Legette et al., 2021; Tomlinson et al., 2003)	4. Academic needs and growth are valued over behavioral prevention. (Tomlinson et al., 2003)
5. Students with similar demographic backgrounds often comprise the same group. (Borman & Dowling, 2010; Condrón, 2007; Van Houtte et al., 2013)	5. Students routinely work with academically, racially, linguistically, and socioeconomically diverse peers. (Tomlinson et al., 1997; Tomlinson et al., 2003)
6. The teacher assumes that students' performance will be stable across all subjects. (Tomlinson et al., 2003)	6. The teacher expects that students' performance will vary across subjects and topics. (Missett et al., 2014; Tomlinson, 2000; Tomlinson et al., 2003).

FIXED GROUPING CHARACTERISTICS

Certain characteristics arise in classrooms that primarily utilize fixed ability grouping. Fixed ability groups are often formed using results from school-mandated, standardized assessments (Haller & Davis, 1981; Tomlinson, 2000). Teachers may consider students' assessment scores from the previous school year or beginning-of-year or end-of-semester diagnostic scores and categorize them as high-achieving, average, or low-achieving based on their results. Fitzgerald et al. (2021) found that many teachers relied on standardized tests to inform grouping decisions, and results from those tests frequently led them to hold fixed conceptions of students' abilities. Thus, once teachers categorize a child as low achieving, per their test score, they may be more inclined to perceive that child as “low ability” and keep them in a lower-level group. Relatedly, fixed groups experience little to no movement throughout the year (Missett et al., 2014; Tomlinson et al., 1997). This means that students designated as cognitively gifted or high achieving may remain in the advanced group for most of the year, while students with learning disabilities, learning challenges, perceived behavioral difficulties, or other reasons (e.g., discriminatory ideas about who is worthy of receiving academic challenge) remain mostly in lower-classified groups (Eder, 1981; Kalogrides & Loeb, 2013; Tyson, 2011). Classifying children in the same level of ability group across multiple subjects (i.e., average for both math and reading) can also be associated with fixed grouping practices if the classification is based more on teachers' personal perceptions and biases than data that supports the grouping (Tomlinson et al., 2003).

Some fixed ability groups are more contingent on students' behavior—and teachers' racialized ideas about behavior—than exhibited ability (Legette et al., 2021; Tomlinson et al., 2003). Since the groups remain relatively consistent throughout the year, the teacher may prioritize having a perceived “manageable” blend of students within each group. Eder (1981) discovered that teachers were more likely to group students based on their behavior, placing children with higher levels of distractibility and interruption into lower groups, regardless of their current readiness levels. This resulted in those groups covering much less content within the allotted instructional time (than the more behaved groups), which perpetuated a self-fulfilling prophecy of “low-ability” groups learning and achieving

less. Kim et al. (2020) also noted that elementary teachers most prioritized behavior in their grouping decisions, specifically aiming to prevent combinations of classmates who could generate behavioral issues when paired—even if they displayed similar levels of readiness for the content. Furthermore, in fixed settings, it is more likely that students with shared demographic characteristics (i.e., racial, linguistic, socioeconomic) will be placed in similar groups (Borman & Dowling, 2010; Condrón, 2007; Van Houtte et al., 2013). Teachers often hold deficit mindsets, perceiving children from lower socioeconomic, single-family, racial minority, and/or immigrant homes as having parents less invested in education or having less exposure to early learning experiences; this leads them to more frequently place and keep these children in lower-ability groups (Calarco, 2014; Gordon & Nocon, 2008; Van Houtte et al., 2013). Put together, fixed grouping typically perpetuates inequitable outcomes for historically marginalized students.

FLEXIBLE GROUPING CHARACTERISTICS

Instead of a reliance on school-mandated, standardized tests, flexible grouping typically relies upon frequent and formative assessments conducted by the teacher (Borland et al., 2002). The teacher might conduct pre-assessments aligned with the curriculum prior to new units, so they can determine students' readiness for that particular subject and topic (Tomlinson et al., 2003). Exit tickets, which are brief quizzes reviewing the major concepts of the lesson, also inform the teacher of students' daily mastery, so they can evaluate if the child needs additional or different support for the following lesson. As they evaluate students' pre-assessment and exit tickets responses, the teacher considers students' relative strengths and weaknesses and flexibly moves them to the group with which they share the most commonalities at that time. In flexible grouping, teachers can quickly respond to students' changes in needs, achievement levels, and motivations by moving them to the group that will best serve their present status.

For example, if a teacher is working with a grade-level group on a fractions unit and notices that one student continues to speed through learning tasks, answering all questions correctly, they may move the student into the advanced-level group the following day to explore deeper fractional concepts with open-ended application opportunities. For the next unit, the teacher would provide another pre-assessment related to that unit's content to determine the best-fit group for that student again; after all, excelling in fractions does not necessarily guarantee advanced readiness for the following unit's concepts. Such a model requires deliberate teacher planning, not only in structuring the learning time and managing student movement, but also in having frequent, formative assessments readily available (Castle et al., 2005; Rubenstein et al., 2015). A strategic and respectful means of communicating the assignment of groups is required; teachers might use something like a digital/written three-by-three table chart with group names, students' names, and a rotation schedule attached that can easily be updated to accommodate the flexibility of student movement and differentiated tasks as need (Tomlinson et al., 2003).

Missett et al. (2014) described how students sorted into seemingly homogeneous groups (e.g., students with learning disabilities, students at grade level, and students identified as gifted or high-achieving) remain inherently heterogeneous even from each other; two identified gifted children (who are often assumed to be similar based on the label) will not hold the same strengths as each other and may exhibit different learning

and motivational struggles (Tomlinson et al., 2003). When educators lock students into one ability level across subjects, they discount the potential strengths and needs those children hold and carry into learning (Tomlinson, 2000; Tomlinson et al., 2003). Thus, flexible grouping assumes that children will perform differently across subjects and topics. It also discourages sorting students based on perceived behavior and, instead, encourages (a) considering students' strengths, interests, and learning motivations related to the concept/skill of focus and (b) pairing them in groups with peers who might complement those well (Kim et al., 2020). It is a higher priority for academic needs to be met in the moment than it is to prevent behaviors from occurring (Tomlinson et al., 2003). This means that if two students chat frequently or routinely disagree but share similar learning patterns for the present unit, the teacher will still place them in the same group—and use alternative management strategies to keep them focused on learning. Since students can move flexibly across groups, it is more common to see peers across diverse demographics and educational backgrounds interacting with and learning from each other (Tomlinson et al., 1997; Tomlinson et al., 2003). In sum, flexible grouping practices aim to ensure that all children learn in a best-fitting environment throughout the scope of the school year, and the teacher utilizing the practices assumes that those environments will regularly change across time, subject, and topic.

EFFECTS OF FLEXIBLE GROUPING AND FIXED GROUPING

In most classrooms, teachers likely adopt a variety of both kinds of grouping practices, based on their current beliefs, capacities, and resources. There are several reasons why researchers endorse the use of flexible grouping over fixed grouping, though. Fixed grouping perpetuates the long-standing conception that ability is inherent and stable over time—someone either is or is not “high-achieving.” This norm has been confronted by research regarding the malleable, flexible nature of intelligence and the impact of environment on one's development (Barab & Plucker, 2002). However, the societal messaging around stratifications (e.g., those who belong in “low” or “high” strata deserve to be placed in them based on their displayed effort; people are born with a certain level of ability and cannot change) reinforces fixed ability thinking as truth (Gamoran, 1989; von Hippel et al., 2018).

Students in relatively fixed ability groups may then be susceptible to fixed ability thinking, where they internalize societal messaging that they are either inherently good or bad at a given domain, based on how they are repeatedly sorted (Fitzgerald et al., 2021; Hargreaves, 2019; Marks, 2013; Tyson, 2011). They may also experience stereotype threat, which describes how individuals' performance suffers from awareness that the identity group(s) to which they belong are not expected to do well (Hartley & Sutton, 2013). Fixed ability thinking acts as a barrier for students in acquiring the motivation and efficacy needed to persist when they face learning challenges (Marks, 2013). Stereotype threat can activate students' anxiety and ultimate disengagement from learning to protect themselves from feared future failures (Steele & Aronson, 1995). Although most teachers do not explicitly articulate each groups' level or designation, students may be able to infer it, based on their personal grades or comparisons to other peer groups, as witnessed in Hargreaves's (2019) study of elementary students who compared their fixed ability groups and test scores with others and subsequently altered their expectations of their personal

abilities. Since students are moved more frequently with flexible grouping and can also be sorted based on interests and strengths, it is less likely that students will attach themselves and their academic self-concept to one level of ability (Marks, 2013).

Perhaps the greatest criticism of fixed grouping practices is that when used repeatedly and primarily as the mode of sorting students, they can perpetuate historical cycles of inequality in schools (Borland et al., 2002; Buttaro et al., 2010; Condrón, 2007; Eder, 1981; Plucker & Peters, 2016). Students of color, students from lower socioeconomic groups, students with disabilities, males, and students from single-parent families are frequently overrepresented in low-ability groups, while White, female, and middle- or upper-middle class students predominantly comprise high-ability groups (Adodo & Agbayewa, 2011; Becton, 2018; Condrón, 2007; Ford, 2011). This can result in historically marginalized groups receiving less access to rigorous instruction and less exposure to classmates who may enrich their thinking (Adodo & Agbayewa, 2011; Kalogrides & Loeb, 2013).

Flexible grouping, on the other hand, has been suggested to provide benefits to both students and teachers (Castle et al., 2005; Plucker & Peters, 2016; Rubenstein et al., 2015; Slavin, 1987). By measuring significant differences in achievement score percent changes over time, Castle et al. (2005) found that in a high-needs school employing flexible grouping over five years, increases in the percentages of students scoring at mastery increased from 10% to 57%. In a study of third grade teachers who were given a pre-differentiated curriculum with included pre-assessments and tiered learning activities (to ease the planning load for teachers), Rubenstein et al. (2015) confirmed that students enjoyed their flexible groups, and their teachers reported increased student engagement with both academic content and peers. This makes sense because flexible movement grants them frequent exposure to different kinds of learners and thinking, which may also support expansion of social networks and a greater sense of community in the classroom. Furthermore, Carol Tomlinson, perhaps one of the most prolific scholars of DI, consistently touts flexible grouping as the most appropriate and respectful means to meet diverse students' learning needs and ensure that all students, regardless of entry point, learn in the classroom (Tomlinson, 2000; Tomlinson et al., 1997; Tomlinson et al., 2003).

Although flexible grouping can address several equity concerns associated with ability grouping, it poses a few practical challenges. Borland et al. (2002) reasonably argued that it is easier to advocate for flexible grouping than it is to implement it because it (a) presents scheduling challenges for administrators, (b) constitutes a change in school culture to adopt curricular differentiation practices, and (c) requires more instructional resources and time for teachers to implement it well. Furthermore, teachers can believe they use a model of flexible grouping, when, in reality, they do not differentiate tasks, materials, or content to the standard of Tomlinson's (2000) DI framework (Chandra Handa, 2020; Maker & Schiever, 2005; Missett et al., 2014; Pozas et al., 2020). When DI is not accomplished, the positive learning effects associated with ability grouping do not occur (Slavin, 1987). Similarly, some teachers do not use formative assessments correctly and may group students based on their perceptions (which increases the potential for bias) instead of what the assessments reveal; this is the primary risk for continued inequity via flexible grouping (Missett et al., 2014). For example, if a fourth-grade emergent bilingual student scores highly on a division pre-assessment, thus constituting the need for advanced differentiation, but the teacher assumes the child may not be able to access the content due to linguistic barriers, then the teacher may place the child in a lower group,

regardless of their pre-assessment results. This would prevent that child from receiving learning opportunities fit to their level of readiness, based on the teacher's assumptions or bias.

Each of these criticisms are valid, proving that flexible grouping practices are not a fix-all on their own; the model relies upon teachers garnering pedagogical-psychological knowledge around how to facilitate groups, how to reflect upon their own biases, and how to identify appropriate times to move children across groups (Heyder et al., 2017; Santangelo & Tomlinson, 2012). Unfortunately, several studies show that this is not a prioritized knowledge base in teacher preservice programs or in-service teacher professional development (Brigandi et al., 2019; De Neve et al., 2015; Evans & Waring, 2008; Fitzgerald et al., 2021; Rubenstein et al., 2015; Santangelo & Tomlinson, 2012). Therefore, certain conditions at the training and K-12 school levels are needed to shift teachers' grouping practices from more fixed in nature to more flexible.

ORGANIZATIONAL INFLUENCES ON ABILITY GROUPING

In understanding systemic ability grouping patterns in American schools, it is helpful to consider teacher training programs and school organizations acting as influential forces in ability grouping, rather than just examining individual teachers' practices (Buttaro et al., 2010). Tomlinson et al. (1994) attribute fixed grouping practices to teacher training programs that insufficiently cover grouping pedagogy with teacher candidates. She argues that when the programs do not provide sufficient modeling—either in their own instruction of teachers or in practicum teaching opportunities—teachers do not have a comprehensive model to which they can refer for their own grouping practices. Once they are placed in a classroom, expected to manage complex responsibilities with little preparation for grouping or the needed capacity for decision-making required of it, it makes sense that they rely upon more traditional, fixed notions of ability, as likely witnessed in their own schooling experiences and as modeled by their school leadership.

Buttaro et al. (2010) attribute observed de facto segregation via ability grouping to three organizational features of schools: structural, cultural, and political. The structural aspect of school organizations relates to the school's characteristics, such as their resources, school size, class size, leadership, and school body demographics. Looking at kindergarten ability grouping data in a large, nationally representative dataset, the authors found that schools with higher proportions of students of color were more likely to utilize within-class ability grouping; they used a composite variable to determine percent minority population versus percent White, so they did not determine any further racial differences. They also discovered that teachers with larger class sizes more frequently used within-class ability grouping to support the large spread of student needs; their data did not differentiate whether grouping was fixed or flexible.

Furthermore, school organizations often do not supply teachers with the curricular materials needed for effective flexible grouping (Harris, 2010). As mentioned earlier, flexible grouping relies upon frequent, formative assessment, which means that teachers readily require unit pre-assessments and lesson exit tickets (Tomlinson et al., 2003). Considering that many elementary teachers must plan daily lessons for all subjects, and then add small group lessons that require different types of instruction for some of those subjects, they have much to plan, often with insufficient planning time provided by school

leaders (Wu, 2013). If the school's mandated curriculum does not supply pre-made pre-assessments and formative assessments, teachers must create them on their own, which is unlikely to occur if they have not been trained in it or do not understand the value behind it. It is more common for district curriculums to be standardized in nature, centered around grade-level content, with little resources for assessing academically diverse students or differentiating content for them (Tomlinson, 2000). In prioritizing planning of instruction, it is understandable that teachers may not have time to create such formative assessments and therefore rely upon standardized assessments' designated ability levels to group students and standardized curriculum to teach groups.

The cultural feature of school organizations represents the norms and ideas most valued by stakeholders within the school (Buttaro et al., 2010). If school administration highly values certain behaviors or exudes fixed ability thinking about students or teachers, it is likely that those values become normed within the school. Teachers' personal backgrounds and beliefs or the school's specific values may inform how they define ability, which leaves any child not in alignment with what is considered "able" vulnerable to inappropriate group placement (Bradbury, 2018). For example, if a teacher values students who complete and submit homework, they may consider a child who consistently puts forth effort on homework as being better suited to a high-ability group than a disorganized student (Van Houtte et al., 2013). Often, school leaders and teachers are more likely to value students whose academic, behavioral, and social skills align with their own background, so their cultural norms shape how they perceive students' ability (Condron, 2007). Condron noted how around 80% of the teaching workforce is comprised of White women, so they may be more likely to unevenly distribute rewards and higher placements towards White students, which could explain the racial disparities often seen in ability groups. Borman and Dowling (2010) similarly found that teachers were more favorably biased towards middle-class students in their grouping decisions, sorting them into groups provided with more enriching tasks.

Finally, political aspects of school organizations may place pressure upon teachers to more frequently utilize fixed practices. For instance, school organizations and site leadership highly concerned with standardized test scores may encourage teachers and staff to practice educational triage through small group instructional time (Bradbury, 2018). Educational triage occurs when students are sorted into three ability groups: safe (meaning they will likely pass the state's standardized test); borderline (meaning they may be close to passing); or hopeless (meaning they are deemed too behind to pass grade-level questions within the year) (Booher-Jennings, 2005; Bradbury, 2018; Tomlinson, 2000). Once students are grouped, these schools ration and allot resources (such as length of instructional small group time, amount of instructional coaching and intervention time, and tutoring opportunities) mostly towards students identified as borderline to increase their likelihood of passing. This rationing—distributed to children in relatively fixed groups—blocks access to instructional resources for students within the two other groups. It is inherently unequal, yet the political pressure imposed upon teachers to improve test scores informs their grouping practices (Bradbury, 2018). Furthermore, if certain parents hold political clout within the school organization (e.g., involvement in Parent-Teacher-Association, frequent volunteering, frequent fundraising donations), they may be more able to assert pressure on school leaders and teachers, who in turn might reward their child with higher placements in groups, access to enrichment opportunities, or

designations of giftedness (Gordon & Nocon, 2008; Harris, 2010; Oakes et al., 1997). When parents frequently communicate with the teacher, asking for help and extension activities to conduct at home, the teacher might conflate effort with ability, thus perceiving the child to have higher abilities than actually exhibited (Calarco, 2014; Gordon & Nocon, 2008). All these structural, cultural, and political features work in tandem and can drive the extent of fixedness or flexibility in teachers' practices.

NEEDED CONDITIONS FOR SYSTEMIC IMPROVEMENT

Teacher training programs and school organizations should consider solutions targeted towards the outlined organizational features and their related issues that drive inequitable grouping in elementary classrooms. Suggestions for consideration are presented below.

PRESERVICE CONDITIONS

Concentrated efforts should be implemented in preservice training programs to better equip teacher candidates to understand and manage a diverse body of learners. De Neve et al. (2015) called for thorough modeling and subsequent explanations of decision-making by course professors, so preservice teachers understand what drives grouping decision-making and can visualize how it is organized and implemented in the classroom. They also recommended that preservice teachers conduct lessons in a lab classroom with each other, where the course professor and classmates can provide subsequent feedback and strategies for future consideration. Because teachers' biases can shape how they view students' abilities or potential, preservice programs could also incorporate critical reflections on personal biases and how they can manifest in teachers' mindsets around ability and grouping of students (Evans & Waring, 2008). Relatedly, Ford (2011) urged programs to train teachers in multicultural giftedness, so they become aware of how potential and talent manifest differently across cultures and environmental contexts and can appreciate talent in children with backgrounds different than their own—an important notion, considering the demographics of the largely White and female teaching workforce.

For flexible grouping to be effective, per Slavin's (1987) findings, teachers should also understand the pedagogy of DI. As mentioned earlier, teachers often believe they differentiate their instruction per small group, but several studies have shown a lack of understanding in what qualifies as true DI of content, process, and product (Maker & Schiever, 2005; Missett et al., 2014; Pozas et al., 2020; Santangelo & Tomlinson, 2012; Tomlinson et al., 1994). Therefore, researchers recommend preservice programs dedicate at least one course to DI, specifically highlighting how to use pre-assessments and other formative assessments to inform equitable and effective grouping decisions (Fitzgerald et al., 2021; Haller & Davis, 1981; Tomlinson et al., 1994). In such classes, leaders could train teacher candidates to recognize academic readiness and not conflate it with compliance; teachers should understand that perceived positive behavior does not translate to qualification for a certain level of ability group (Tomlinson et al., 1994).

Since DI can be difficult to facilitate, teachers should understand how to manage a flexibly grouped, DI classroom and practice leading flexible ability groups during student teaching experiences (Tomlinson et al., 2003). For example, programs can teach candidates how to respectfully communicate and change group assignments, as well as

how to establish behavioral and work expectations for independent or partner work time. Teachers also should understand how to develop appropriate, engaging work materials for these times, as well as how to differentiate tasks and materials (via scaffolding or enrichment) for their small group instruction. If these management items are addressed prior to entry in the field, teachers might be better equipped to flexibly group students and enjoy higher amounts of efficacy in DI (Dixon et al., 2014), which in turn should encourage continued use of flexible grouping practices (Poulou et al., 2018). Even with such training, however, it can take several years to acquire and seamlessly employ these skills (Santangelo & Tomlinson, 2012; Tomlinson et al., 1994).

ORGANIZATIONAL CONDITIONS

In addition to teacher training, in-service conditions must also be satisfied for flexible grouping to develop, especially for most veteran teachers who have not been exposed to the recommended pre-service training. Brigandi et al. (2019), in following teachers' longitudinal development of DI and grouping practices, observed how professional development (PD) increased teachers' grouping knowledge and DI toolkits, but veteran teachers' beliefs and attitudes about grouping were difficult to transform. Therefore, they urged school administrators to avoid one-day PD seminars that might not inspire immediate change and instead supply meaningful, ongoing supports with trusted colleagues that will foster experimentation and risk-taking with flexible grouping and DI. For example, they described how sustained coaching with an instructional coach who models flexible grouping and DI practices can motivate a teacher to try it in their classroom. Relatedly, if teachers are to attempt challenging new practices, school administrators must grant them autonomy to experiment with different methods and provide grace when they fail (De Neve et al., 2015). Professional learning communities (PLCs), which are comprised of teachers committed to studying, sharing, and practicing strategies around a shared topic, are another motivating tool for teacher transformation because they are teacher-led and practical in nature. De Neve et al. (2015) and Tomlinson et al. (2003) highlighted how PLCs around flexible grouping and DI can build a stronger sense of self-efficacy for such complex tasks, and they allow teachers to glean new strategies practiced and approved by colleagues.

Since flexible grouping requires additional instructional and assessment materials, district organizations could structurally ease the planning burden by purchasing high-quality pre-differentiated curricula that includes all needed formative assessments for grouping purposes and tiered activities for all groups' levels of readiness (Azano et al., 2011; Callahan et al., 2015; McCoach et al., 2014; Plucker & Peters, 2016; Rubenstein et al., 2015). High-quality, pre-differentiated curricula can raise teachers' awareness of diverse student needs and support them in adopting flexible grouping practices and providing appropriate lessons for each group (Rubenstein et al., 2015). Thus, if organizations met this structural need, they could reduce the planning and instructional load that flexible grouping can place upon teachers.

Cultural shifts within school organizations also must occur for flexible grouping practices to become more commonly used than fixed ability grouping practices (Harris, 2010). De Neve et al. (2015) suggested that such cultural shifts require support among colleagues and a shared sense of responsibility for students across the building, so administrators

can facilitate staff conversations where such values are communicated and considered. Administrators seeking to increase flexibility in teachers' grouping practices might also consider how their personal practices have fueled fixed ability thinking in their schools and then consider what steps must be taken in staff meetings, PD sessions, informal conversations, and teacher evaluation meetings to alter the present school culture.

To address political pressure from influential parents seeking high grouping placement for their children, administrators can communicate to parents the school's commitment to flexible grouping at the beginning of the school year, providing empirical justification that outline its benefits, so expectations are established about how small group instruction will function. Then, if a parent places pressure on a teacher about their child's "level" or placement in a group, the administrator can provide support to the teacher and offer to sit in on the parent meeting, if needed. Further, Tomlinson et al. (2003) insisted that schools shift their rigid concerns around communal testing into more concern for students' individual strengths and needs; this would remove the political stress from teachers to isolate students into groups and grant more freedom to use formative assessments and experiment flexibly in their classrooms.

POTENTIAL CHALLENGES TO CONSIDER IN THE PURSUIT OF EQUITABLE GROUPING

Structural racism and classism are embedded into school funding formulas based on local tax revenue and thus result in schools with higher proportions of students of color and students from low-income backgrounds receiving less dollars per pupil (Baker et al., 2020; Sosina & Weathers, 2019; Weathers & Sosina, 2022). This reality may make some of the above-mentioned recommendations challenging to implement in historically marginalized communities. For example, schools with less funds may not be able to afford structural changes like purchasing a pre-differentiated curriculum or providing ongoing coaching dedicated to grouping and DI. In these cases, school leaders might consider shifting mandatory staff meetings or PD time into time in which grade level teams can collaboratively plan and design pre-assessments and formative assessments together. Collaborative planning eases the practical and cognitive load on teachers (Thousand et al., 2006), so if they are granted time to design assessments aligned to their current curriculum, that might support their use of flexible grouping. Furthermore, instead of providing one-off PDs for the year, as is traditionally done in public school settings, the principal might instead encourage teachers to conduct individual or collaborative action research related to their grouping and DI practices for the school year; typical time allotted for PD could be granted to teachers to reflect on their goals, use their collected data to monitor their progress, and revise goals and plans for action as they go (Mitchell et al., 2009). Innovation may be required to better equip teachers to flexibly group when funds are limited, but it is possible and imperative to prioritize flexible grouping development in divested communities.

Furthermore, schools with higher proportions of students of color and low-income students often have more difficulty recruiting and retaining high-quality teachers, with structural racism and classism similarly driving these effects (Cherng et al., 2022; Lane et al., 2018). This could result in the students who most require equitable within-class small group instruction not accessing teachers with the pedagogical-psychological knowledge

to facilitate it. If administrators in schools with higher proportions of historically marginalized students notice that their teaching body does not hold pedagogical-psychological knowledge for equitable grouping, then it is critical that they prioritize training efforts, particularly citing the evidence of how grouping practices have historically harmed students and how flexible grouping can prevent those harms from repeating.

CONCLUSION

Elementary teachers undoubtedly carry a great load as they seek to meet students' increasingly heterogeneous academic, social, and emotional needs (Tomlinson et al., 1994). To simplify their work, they often group students into more homogeneous ability groups, so they can teach students according to their needs (Slavin, 1987; Sørensen, 1970). Such ability groups can take on various combinations of fixed and flexible characteristics, meaning that in different settings, students can sometimes become locked into one ability group or can move freely across groups fit to current levels of readiness. While there are multiple reasons to explain why certain teachers sometimes utilize more fixed ability grouping practices, it is important to consider how such practices impact students, especially those who have been historically marginalized through sorting and grouping patterns in schools (Borland et al., 2002; Buttaro et al., 2010; Ford, 2011). Teachers' grouping practices are shaped by school organizational features and their own personal factors (such as how they perceive ability), which often works together to perpetuate longstanding, inequitable grouping of certain groups (Tomlinson et al., 1994). Unless teacher training programs and school organizations employ targeted efforts to shift their organizations' values and resources towards flexible ability grouping practices, teachers will likely remain dependent on more fixed practices since they are professionally familiar and easier to implement. Given the generational cycles of inequity hosted in American schools, that is something we can no longer accept.

REFERENCES

- Adelson, J. L., & Carpenter, B. D. (2011). Grouping for achievement gains: For whom does achievement grouping increase kindergarten reading growth? *Gifted Child Quarterly*, 55(4), 265–278. <https://doi.org/10.1177/0016986211417306>
- Adodo, S. O., & Agbayewa, J. O. (2011). Effect of homogeneous and heterogeneous ability grouping class teaching on a student's interest, attitude, and achievement in integrated science. *International Journal of Psychology and Counseling*, 3(3), 48–54. <https://doi.org/10.5897/IJPC.9000013>
- Alexander, K. L., & Cook, M. A. (1982). Curricula and coursework: A surprise ending to a familiar story. *American Sociological Review*, 47(5), 626–640. <https://doi.org/10.2307/2095163>
- Allington, R. L. (1983). The reading instruction provided readers of differing reading abilities. *The Elementary School Journal*, 83, 548–559. <https://doi.org/10.1086/461333>
- Altintas, E., & Ozdemir, A. S. (2015). Evaluating a newly developed differentiation approach in terms of student achievement and teachers' opinions. *Educational Sciences: Theory and Practice*, 15(4), 1103–1118. <https://doi.org/10.12738/estp.2015.4.2540>
- Azano, A., Missett, T. C., Callahan, C. M., Oh, S., Brunner, M., Foster, L. H., & Moon, T. R. (2011). Exploring the relationship between fidelity of implementation and academic achievement in a third-grade gifted curriculum: A mixed-methods study. *Journal of Advanced Academics*, 22(5), 693–719. <https://doi.org/10.1177/1932202X11424878>
- Baker, B. D., Srikanth, A., Cotto, R., & Green, P. C. (2020). School funding disparities and the plight of Latinx children. *Education Policy Analysis Archives*, 28(135). <https://doi.org/10.14507/epaa.28.5282>
- Barab, S. A., & Plucker, J. A. (2002). Smart people or smart contexts? Talent development in an age of situated approaches to learning and thinking. *Educational Psychologist*, 37, 165–182. https://doi.org/10.1207/S15326985EP3703_3
- Barr, R., & Dreeben, R. (1983). *How Schools Work*. University of Chicago Press.
- Becton, V. (2018). Re-examining the resurgence of ability grouping in public schools: Does it really impact academic achievement and the social and emotional development of students with disabilities? *International Journal of Arts and Sciences*, 11(1), 253–268.
- Booher-Jennings, J. (2005). Below the bubble: “Educational triage” and the Texas accountability system. *American Educational Research Journal*, 42(2), 231–268. <https://doi.org/10.3102/00028312042002231>
- Borland, J. H., Horton, D., Subotnik, R. F., Shiang-Jiun, C., Miran, C., Freeman, C., & Yu, J. (2002). Ability-grouping and acceleration of gifted students: Articles from the Roeper Review. *Roeper Review*, 24, 100–102. <https://doi.org/10.1080/02783190209554138>

- Borman, G. D., & Dowling, M. (2010). Schools and inequality: A multilevel analysis of Coleman's equality of educational opportunity data. *Teachers College Record*, 112(5), 1201-1246. <https://doi.org/10.1177/016146811011200507>
- Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. K. Brown (Ed.), *Knowledge, education, and cultural change: Papers in the sociology of education* (pp. 71-112). Tavistock Publications.
- Bradbury, A. (2018). The impact of the phonics screening check on grouping by ability: A 'necessary evil' amid the policy storm. *British Educational Research Journal*, 44(4), 539-556. <https://doi.org/10.1002/berj.3449>
- Brigandi, C. B., Gilson, C. M., & Miller, M. (2019). Professional development and differentiated instruction in an elementary school pullout program: A gifted education case study. *Journal for the Education of the Gifted*, 42(4), 362-395. <https://doi.org/10.1177/0162353219874418>
- Brookwood, M. (2021). *The orphans of Davenport: Eugenics, the Great Depression, and the war over children's intelligence*. Liveright.
- Buttaro A., Jr., Catsambis, S., Mulkey, L., & Steelman, L. C. (2010). An organizational perspective on the origins of instructional segregation: School composition and use of within-class ability grouping in American kindergartens. *Teachers College Record*, 112(5), 1300-1336. <https://doi.org/10.1177/016146811011200503>
- Calarco, J. M. (2014). Coached for the classroom: Parents' cultural transmission and children's reproduction of educational inequalities. *American Sociological Review*, 79, 1015-1037. <https://doi.org/10.1177/00031224114546931>
- Callahan, C. M., Moon, T. R., Oh, S., Azano, A. P., & Hailey, E. P. (2015). What works in gifted education: Documenting the effects of an integrated curricular/instructional model for gifted students. *American Educational Research Journal*, 52(1), 137-167. <https://doi.org/10.3102/0002831214549448>
- Castle, S., Deniz, C. B., & Tortora, M. (2005). Flexible grouping and student learning in a high-needs school. *Education and Urban Society*, 37(2), 139-150. <https://doi.org/10.1177/0013124504270787>
- Chandra Handa, M. (2020). Examining students' and teachers' perceptions of differentiated practices, student engagement, and teacher qualities. *Journal of Advanced Academics*, 31(4), 530-568. <https://doi.org/10.1177/1932202X20931457>
- Cherng, H. S., Halpin, P. F., & Rodriguez, L. A. (2022). Teaching bias? Relations between teaching quality and classroom demographic composition. *American Journal of Education*, 128(2), 171-201. <https://doi.org/10.1086/717676>
- Condon, D. J. (2007). Stratification and educational sorting: Explaining ascriptive inequalities in early childhood reading group placement. *Social Problems*, 54(1), 139-160. <https://doi.org/10.1525/sp.2007.54.1.139>

- Corbett Burris, C., Welner, K., & Murphy, J. (2008). Accountability, rigor, and detracking: Achievement effects of embracing a challenging curriculum as a universal good for all students. *Teachers College Record*, 110(3), 571-607. <https://doi.org/10.1177/016146810811000301>
- De Neve, D., Devos, G., & Tuytens, M. (2015). The importance of job resources and self-efficacy for beginning teachers' professional learning in differentiated instruction. *Teaching and Teacher Education*, 47, 30-41. <https://doi.org/10.1016/j.tate.2014.12.003>
- Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, 37(2), 111-127. <https://doi.org/10.1177/0162353214529042>
- Eder, D. (1981). Ability grouping as a self-fulfilling prophecy. *Sociology of Education*, 54, 151-162. <https://doi.org/10.2307/2112327>
- Ellison, B. J., & Hallinan, M. T. (2004). Ability grouping in Catholic and public schools. *Catholic Education: A Journey of Inquiry and Practice*, 8(1), 107-129. <https://ejournals.bc.edu/index.php/cej/article/view/571>
- Esposito, D. (1973). Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. *Review of Educational Research*, 43(2), 163-180. <https://doi.org/10.3102/00346543043002163>
- Evans, C., & Waring, M. (2008). Trainee teachers' cognitive styles and notions of differentiation. *Education and Training*, 50(2), 140-154. <https://doi.org/10.1108/00400910810862128>
- Fitzgerald, L., Hunter, J., & Hunter, R. (2021). Shifting teacher practices in relation to grouping: Gap gazing or strengths focused approaches. *Mathematics Teacher Education & Development*, 23(3), 97-110. <https://files.eric.ed.gov/fulltext/EJ1321057.pdf>
- Ford, D. Y. (2011). *Multicultural gifted education (2nd ed.)*. Prufrock Press.
- Gambrell, L., Wilson, R. M., & Walter, N. (1981). Classroom observations of task-attending behaviors of good and poor readers. *Journal of Educational Research*, 74, 400-404.
- Gamoran, A. (1986). Instructional and institutional effects of ability grouping. *Sociology of Education*, 59, 185-199. <https://doi.org/10.2307/2112346>
- Gamoran, A. (1989). Rank, performance, and mobility in elementary school grouping. *The Sociological Quarterly*, 30(1), 109-123. <https://doi.org/10.1111/j.1533-8525.1989.tb01514.x>
- Goldberg, M. L., Passow, A. H., & Justman, J. (1966). *The effects of ability grouping*. Teachers College Press.
- Good, T. L., & Brophy, J. E. (1972). Behavioral expression of teacher attitudes. *Journal of Educational Psychology*, 63, 617-624. <https://doi.org/10.1037/h0034069>
- Gordon, V., & Nocon, H. (2008). Reproducing segregation: Parent involvement, diversity, and school governance. *Journal of Latinos and Education*, 7(4), 320-339. <https://doi.org/10.1080/15348430802143634>

- Haller, E. J., & Davis, S. A. (1981). Teacher perceptions, parental social status and grouping for reading instruction. *Sociology of Education*, 54, 162–174. <https://doi.org/10.2307/2112328>
- Hargreaves, E. (2019). Feeling less than other people: Attainment scores as symbols of children's worth. *Forum for Comprehensive Education*, 16(1), 53–66. <http://doi.org/10.15730/forum.2019.61.1.53>
- Harris, D. M. (2010). Curriculum differentiation and comprehensive school reform: Challenges in providing educational opportunity. *Educational Policy*, 25(5), 844–884. <https://doi.org/10.1177/0895904810386600>
- Hartley, B. L., & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development*, 84(5), 1716–1733. <https://doi.org/10.1111/cdev.12079>
- Heyder, A., Bergold, S., & Steinmayr, R. (2017). Teachers' knowledge about intellectual giftedness: A first look at levels and correlates. *Psychology and Learning*, 17(1), 27–44. <https://doi.org/10.1177/1475725717725493>
- Ireson, J., & Hallam, S. (1999). Raising standards: Is ability grouping the answer? *Oxford Review of Education*, 25(3), 343–358. <https://doi.org/10.1080/030549899104026>
- Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher*, 42(6), 304–316. <https://doi.org/10.3102/0013189X13495087>
- Kim, S., Lin, T. J., Chen, J., Logan, J., Purtell, K. M., & Justice, L. M. (2020). Influence of teachers' grouping strategies on children's peer social experiences in early elementary classrooms. *Frontiers in Psychology*, 11, 3583. <https://doi.org/10.3389/fpsyg.2020.587170>
- Lane, E., Linden, R., & Stange, K. (2018). *Socioeconomic disparities in school resources: New evidence from within-districts*. University of Michigan. <http://www-personal.umich.edu/~kstange/papers/LaneLindenStangeOct2018.pdf>
- Legette, K. B., Halberstadt, A. J., & Majors, A. T. (2021). Teachers' understanding of racial inequity predicts their perceptions of students' behaviors. *Contemporary Educational Psychology*, 67, 102014. <https://doi.org/10.1016/j.cedpsych.2021.102014>
- Loveless, T. (2009). *Tracking and detracking: High achievers in Massachusetts middle schools*. Thomas B. Fordham Institute.
- Maker, C. J., & Schiever, S. W. (2005). *Teaching models in education of the gifted* (3rd ed.). Pro-Ed.
- Marks, R. (2013). "The blue table means you don't have a clue": The persistence of fixed-ability thinking and practices in primary mathematics in English schools." *Forum: For Promoting 3-19 Comprehensive Education*, 55(1), 31–44. <https://doi.org/10.2304/forum.2013.55.1.31>
- McCoach, D. B., Gubbins, E. J., Foreman, J., Rubenstein, L. D., & Rambo-Hernandez, K.

- E. (2014). Evaluating the efficacy using predifferentiated and enriched mathematics curricula for grade 3 students: A multisite cluster-randomized trial. *Gifted Child Quarterly*, 58(4), 272–286. <https://doi.org/10.1177/0016986214547631>
- Missett, T. C., Brunner, M. M., Callahan, C. M., Moon, T. R., & Azano, A. P. (2014). Exploring teacher beliefs and use of acceleration, ability grouping, and formative assessment. *Journal for the Education of the Gifted*, 37(3), 245–268. <https://doi.org/10.1177/0162353214541326>
- Mitchell, S. N., Reilly, R. C., & Logue, M. E. (2009). Benefits of collaborative action research for the beginning teacher. *Teaching and Teacher Education*, 25(2), 344–349. <https://doi.org/10.1016/j.tate.2008.06.008>
- National Center for Education Statistics (2012). *The nation's report card: 2011*. The Nation's Report Card. <https://nces.ed.gov/nationsreportcard/pdf/main2011/2012457.pdf>
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. Yale University Press.
- Oakes, J. (1992). Can tracking research inform practice? Technical, normative, and political considerations. *Educational Researcher*, 21(4), 4–47. <https://doi.org/10.3102/0013189X021004012>
- Oakes, J., Stuart Wells, A., Jones, M., & Datnow, A. (1997). Detracking: The social construction of ability, cultural politics, and resistance to reform. *Teachers College Record*, 98(3), 483–510. <https://doi.org/10.1177/016146819709800305>
- Plucker, J. A., & Peters, S. J. (2016). *Excellence gaps in education: Expanding opportunities for talented students*. Harvard Education Press.
- Pozas, M., Letzel, V., & Schneider, C. (2020). Teachers and differentiated instruction: Exploring differentiation practices to address student diversity. *Journal of Research in Special Educational Needs*, 20(3), 217–230. <https://doi.org/10.1111/1471-3802.12481>
- Poulou, M. S., Reddy, L. A., & Dudek, C. M. (2018). Relation of teacher self-efficacy and classroom practices: A preliminary investigation. *School Psychology International*, 40(1), 25–48. <https://doi.org/10.1177/0143034318798045>
- Rubenstein, L. D., Gilson, C. M., Bruce-Davis, M. N., & Gubbins, E. J. (2015). Teachers' reactions to pre-differentiated and enriched mathematics curricula. *Journal for the Education of the Gifted*, 38(2), 141–168. <https://doi.org/10.1177/0162353215578280>
- Santangelo, T., & Tomlinson, C. A. (2012). Teacher educators' perceptions and use of differentiated instruction practices: An exploratory investigation. *Action in Teacher Education*, 34, 309–327. <https://doi.org/10.1080/01626620.2012.717032>
- Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of educational research*, 57(3), 293–336. <https://doi.org/10.3102/00346543057003293>
- Sørensen, A. B. (1970). Organizational differentiation of students and educational opportunity. *Sociology of Education*, 43(4), 355–376.

- Sosina, V. E., & Weathers, E. S. (2019). Pathways to inequality: Between-district segregation and racial disparities in school district expenditures. *AERA Open*, 5(3). <https://doi.org/10.1177/2332858419872445>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, 69, 797-811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K-12 students' academic achievement: Findings on two second-order meta-analyses. *Review of Educational Research*, 86(4), 849-899. <https://doi.org/10.3102/0034654316675417>
- Stright, A. D., & Supplee, L. H. (2002). Children's self-regulatory behaviors during teacher-directed, seat-work, and small-group instructional contexts. *The Journal of Educational Research*, 95(4), 235-244. <https://doi.org/10.1080/00220670209596596>
- Thousand, J. S., Villa, R. A., & Nevin, A. I. (2006). The many faces of collaborative planning and teaching. *Theory Into Practice*, 45(3), 239-248. https://doi.org/10.1207/s15430421tip4503_6
- Tomlinson, C. A., Tomchin, E. M., & Callahan, C. M. (1994). Practices of preservice teachers related to gifted and other academically diverse learners. *Gifted Child Quarterly*, 38(3), 106-114. <https://doi.org/10.1177/001698629403800303>
- Tomlinson, C. A., Callahan, C. M., Tomchin, E. M., Eiss, N., Imbeau, M., & Landrum, M. (1997). Becoming architects of communities of learning: Addressing academic diversity in contemporary classrooms. *Exceptional Children*, 63, 269-282. <https://doi.org/10.1177/001440299706300210>
- Tomlinson, C. A. (2000). Reconcilable differences? Standards-based teaching and differentiation. *Educational Leadership*, 58(1), 6-11. <https://www.ascd.org/el/articles/reconcilable-differences-standards-based-teaching-and-differentiation>
- Tomlinson, C. A., Brighton, C., & Hertberg, H. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2/3), 119-145. <https://doi.org/10.1177/016235320302700203>
- Tyack, D. (1974). *The one best system: A history of American urban education*. Harvard University Press.
- Tyson, K. (2011). *Integration interrupted: Tracking, black students, and acting white after Brown*. Oxford University Press.
- Van Houtte, M., Demanet, J., & Stevens, P. A. J. (2013). Curriculum tracking and teacher evaluations of individual students: Selection, adjustment, or labeling? *Social Psychology of Education: An International Journal*, 16(3), 329-352. <https://doi.org/10.1007/s1123-18-013-9216-8>
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general

- pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969. <https://doi.org/10.1037/a0025125>
- von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in reading and math skills forms mainly before Kindergarten: A replication, and partial correction, of “Are schools the great equalizer?” *Sociology of Education*, 91(4), 323–357. <https://doi.org/10.1177/0038040718801760>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Weathers, E. S., & Sosina, V. E. (2022). Separate remains unequal: Contemporary segregation and racial disparities in school district revenue. *American Educational Research Journal*, 59(5), 905–938. <https://doi.org/10.3102/00028312221079297>
- Worthy, J. (2010). Only the names have been changed: Ability grouping revisited. *The Urban Review*, 42(4), 271–295. <https://doi.org/10.1007/s11256-009-0134-1>
- Wu, E. H. (2013). An interview with Carol Tomlinson. *Journal of Advanced Academics*, 24(2), 125–133. <https://doi.org/10.1177/1932202X13483472>