

TO MIX OR NOT TO MIX: A CRITICAL REVIEW OF LITERATURE ON MIXED-AGE GROUPS IN PRIMARY SCHOOLS

Zoe Cronin

University of Cambridge, Cambridge, United Kingdom

znc20@cam.ac.uk

In this review, I explore mixed-age grouping in primary schools, illustrating, through a review of scholarly research, its position within current education paradigms and in the field of education research. I justify my investigation into this topic and explicate my literature search procedure, considering the difficulties around establishing consistent terminology in mixed-age research. I explain various circumstances that give rise to mixed-age groupings and propose using four circumstantial categories – default, by-product, mandate, and preference – as a conceptual framework for understanding mixed-age phenomena. I then summarize findings from methodologically diverse inquiries into the effects of mixed-age grouping. These studies, conducted over the last sixty years, focus on many forms of mixed-age groups from around the world and consider both academic and social outcomes. Broadly speaking, outcome-based findings are inconsistent across time and place. Systematically measured differences are often small or non-existent. In the context of ambiguous empirical findings, I discuss the perspectives held by parents, teachers, and researchers around mixed-age grouping and highlight limitations of utilizing findings from comparative studies to inform education practice. I position the outcomes of my reviewed literature within the proposed circumstantial framework and discuss the implications of this standpoint. I deconstruct arguments for and against mixed-age grouping by posing the question “to mix or not to mix”, offering apparent reasoning for each position. I extend my perspective on the future of mixed-age research, focusing on the need for thorough description and clear definition of all investigated mixed-age groups, and conclude by critically considering mixed-age grouping as a promising education reform.

Keywords: Mixed-age, Monograde, Grouping, Student outcomes, Primary education

Introduction

Student grouping is a foundational and hotly debated principle in education. The ambitious goals of providing “inclusive and equitable quality education” and “lifelong learning opportunities for all” call for efficient and effective organization of human resource (The Global Education 2030 Agenda, 2015, p.1). They also demand the continuous defining and redefining of “quality education” in order to meet the needs of a perpetually evolving society. Grouping structure has implications for both accessibility and quality of education, thus the question of whether to mix age groups in primary schools requires ongoing critical analysis. My professional experiences have allowed me to observe children of all ages in an array of contexts¹. In homes, children typically live with siblings and family members of all ages. In after-school programs, children often span several age years. Public spaces of play and learning (e.g., parks, museums, etc.) welcome children of all ages. Beyond childhood, it is rare to find higher education or professional settings comprised of a single year group of people. Even within the primary school day, it is common practice for all students to have simultaneous break times. Mixed-age grouping is the norm in all these situations aside from one: inside school classrooms. Why is an unnaturally occurring grouping structure mainstream in primary schools? What other groupings are practiced? How do they differ? These questions led me to formally investigate mixed-age grouping in primary schools. In producing this work, I engaged critically with each study I reviewed, considering whether its methodological approach provided valid answers to its questions, whether its conclusions reflected the nature of its findings, and whether its inherent value justified its completion.

An overview of mixed-age literature

To identify key literature, I first searched the term “multiage” on the Education Resources Information Center (ERIC) database, yielding over six-hundred results. Bailey, Werth, Allen, & Sutherland (2016) (listed as the most relevant result) led me to Veenman’s (1995) seminal best-evidence synthesis of findings on the effects of mixed-age grouping in primary schools. From this and several articles in the reference list, I constructed a list of search terms: mixed-age, multiage, multigrade, nongrade, combination class, or vertical group. A search yielded twenty-eight articles with at least one of these search terms in the title from within the last twenty years on the British Education Index (BEI). This literature review is based upon these articles (exclusive of five that examined early childhood or adult education), relevant items from their reference lists, and the exploration of a bit of grey literature (i.e., first several Google results of my aforementioned search terms). Though this is by no means an exhaustive review of literature on this topic, recurring tensions, concepts, authors, and a myriad of methodologies employed around the world seemed to indicate I was approaching saturation. Because I am ultimately interested in working in education in American schools, I limited my outcome-exploration to findings from North America, South America, Europe, and New Zealand. Examination of global mixed-age findings would undoubtedly offer a more complete portrayal

¹ I began working with children in 2006, teaching private dance and piano lessons. I assistant taught at Grinnell College Laboratory Preschool, directed a weekly volunteer program for children ages five to twelve, and was a summer camp Counselor for five years. After receiving a Bachelor’s in Psychology, I did community and school-based social work for several years. Most recently, I implemented a Positive Behaviour Intervention System and play-based learning program while substitute teaching at an elementary school in Oakland, California.

of the holistic mixed-age phenomenon, and is a topic of future interest to me. Notably, a search for the terms “single grade” or “monograde” in a study’s title yielded only six results across both ERIC and BEI, all of which were articles comparing these structures to mixed-age grouping. In my search through databases, reference lists, and grey literature, I did not come across any pedagogical or empirical arguments for single-gradedness as an independent concept (i.e., not defined in relation to mixed-age alternatives). This supports Anderson and Pavan’s (1993) assertion that “there is not, and there has not been, any philosophical or research-based support for continuation of graded structure” (p. xi) and suggests that this absence persists twenty-five years later.

I draw repeatedly from a few key pieces of literature. Slavin (1988) discusses grouping methods historically utilized in primary schools, considering the role of student grouping generally and synthesizing research findings on numerous alternative grouping methods. Slavin investigated mixed-age grouping throughout the 1990s and is frequently cited (Gutierrez & Slavin, 1988; Kallery & Loupidou, 2016; Kelly-Vance, Caster, & Ruane, 2000). Vincent (1999) edited a handbook written for rural educators originally by Bruce Miller, another frequently cited mixed-age researcher (Kelly-Vance et al., Little, 2001; Miller, 1991; Miller, 1993). This work overviews the American history of mixed-age and single-grade education, reviews comparative studies, and considers the role of rural educators. Angela Little, Professor Emerita at the University College London Institute of Education, investigates mixed-age grouping and education reform around the world. In her often-cited 2001 article, she considers mixed-age teaching in an international research and policy context (Aksoy, 2008; Hargreaves, 2001). After more than twenty years teaching and a PhD examining parent perceptions of mixed-age grouping, Linley Cornish became an Associate Professor at the University of New England where she has continued examining mixed-age grouping. Her 2015 article written for curious and concerned parents contemplates learning effects of mixed-age and single-grade classrooms. Recent mixed-age researchers consistently draw from her work (Bailey et al., 2016; Quail & Smyth, 2014; Smit & Engeli, 2015; Smit, Hyry-Bheimmer, & Raggl, 2015).

The generality of “mixed-age grouping” makes it difficult to establish a robust and meaningful terminology. There is a general consensus that in order to qualify as mixed-age (or any of the previously stated search terms) in primary school groupings, the group must include students from more than one grade-level. This simplistic operationalization, however, may conceal critical differences in mixed-age manifestations. It also serves to accentuate the mix of ages that exists within any given grade level. A single grade typically includes children up to a year apart, when even a three-month age difference between children, especially primary school-aged children, can lead to stark physical and developmental differences quite plain to a casual observer. Then again, sometimes a younger child can present as physically, socially, or academically “older” – in the sense in this case of “more mature”. Some researchers bypass these nuances by acknowledging the multiplicity of terms and simply deciding on one to utilize. Stuart, Connor, Cady and Zweifel (2007) discuss the details of a particular “multiage” classroom’s structure without making claims about broader uses of this term. Quail and Smyth (2014) refer exclusively to “multigrade” classes without providing more specificity than the fact that they are comprised of students from more than one grade-level. Huf and Raggl (2015, 2017) and Lindström and Lindahl (2011) state that their observations took place in “age-mixed” and “mixed-age” classes respectively without any insight into how or why these classes were formed. For the remainder of this review, I will refer to groups comprised of students from two or more grade-levels as mixed-age groupings and report

classroom formation and operation details whenever possible. Since Veenman (1995) distinguished between “multigrade” classes formed for “administrative and financial reasons” (p.319) and “multiage” classes formed for “pedagogical and didactic motives” (p.321), many researchers concur that a fundamental difference exists between classes that mix ages out of necessity and those that mix out of choice (Hargreaves, 2001; Huf & Raggl, 2015; Lindström & Lindahl, 2011; Little, 2001; Mason & Burns, 1996). Hargreaves (2001) gives just one example of how reasoning behind the grouping structure could have important implications, stating that in classrooms mixed by choice “there is less likely to be pressure on the teacher to cover a grade-related curriculum” (p.554). Within the context of mixed-age literature, this review aims to acknowledge and explicate the myriad factors, focusing on circumstances, outcomes, and perspectives, that affect the decision whether to mix ages in primary schools.

Circumstances

In an effort to conceptualize the vast mixed-age landscape, I distinguish four categories of mixed-age groups that simply and meaningfully reduce the many circumstances that lead to these groupings in primary schools. Generally speaking, mixed-age groupings are default options, by-products of grouping by another characteristic, mandates from the educational system, or preferences. I utilize the term “circumstance” – a derivative of the Latin *circumstare*, meaning “encircle” or “encompass” – intentionally to reflect the early stage of understanding mixed-age phenomenon. These circumstantial categories are overlapping, rather than rigid.

Default

If there are insufficient numbers of teachers or students, mixed-age grouping becomes the default option in order to ensure access to education. This is often the case in sparsely populated communities (McEwan, 2008; Seban, 2015; Vincent, 1999). Vincent (1999) points out that in rural schools, mixed-age grouping is an unavoidable reality based on economic and geographic necessity. Indeed, up until the rapid urbanization of the early 20th century, one-room schoolhouses were the prevailing school structure presumably for these reasons (Kelly-Vance, Caster, & Ruane, 2000). By the same token, schools or classes geared toward a smaller subset of the student population (e.g., special schools, special education classrooms, gifted programs) are often mixed-age by default (Slavin, 1988).

By-product

Age is but one of numerous characteristics that can be utilized to group students. In “ability-grouping”, for example, children are divided by their perceived or measured ability, often within a particular subject, as in the Joplin Plan that groups students by reading level (Slavin, 1988). Though this important topic is beyond the scope of this review, it is worth noting that “ability-grouping” is steeped in controversy, referred to by Bradbury (2018) as a “necessary evil” (p.1) and by McGilliguddy and Devine (2018) as “an act of symbolic violence” (p.1). In “departmentalization”, the norm in higher education and an increasingly common practice in secondary education, students across age groups divide by subject or discipline (Slavin, 1988). Educators may also divide students by common interest, need for specific skill acquisition, or allow for student self-selection, though these divisions are

typically made within single-grade classrooms as means of creating small groups rather than to divide a school population into classrooms (Hoffman, 2002; Slavin, 1988; Stuart et al., 2007).

Mandate

In some cases, mandates within an educational system require classrooms to mix grade levels in single classrooms, as in Veenman's (1995) "multigrade" classrooms created for "administrative and financial reasons" (p.319). Requirements around class size, perhaps the most common mandated impetus for mixing age groups, might lead to the redistribution of single-grade students into more evenly sized mixed-age classrooms for example (Mason & Burns, 1996; Stuart et al., 2007; Veenman, 1995).

Preference

Mixed-age grouping may be instituted when it is the preferred grouping strategy, and this can manifest in a variety of ways. Nongraded programs, Montessori schools, and Veenman's "multiage" classrooms created for "pedagogical and didactic motives" (p.321) deliberately establish age heterogeneity inside primary classrooms for its perceived educational benefit (Di Lorenzo & Salter, 1965; Huf & Raggl, 2017; Slavin, 1988; Veenman, 1995).

Circumstantial context

Categorization by circumstance serves as a clarifying system with which to comprehend the necessity, prevalence, and intentionality behind mixed-age education. This framework is fluid and overlapping. For example, Anderson and Pavan (1993) describe the "nongraded" method as a comprehensive educational philosophy that requires the pedagogical buy-in of every player in the educational system. Obviously then, mixed-age grouping is a "mandate" of sorts in this structure, in addition to the preferred pedagogy. Though student or teacher numbers in sparsely populated areas may default schools to mixed-age grouping, teachers and administrators may very well believe it is the preferred structure. In fact, default and mandated mixed-age classrooms often decide to embrace mixed-age pedagogy, making virtue out of necessity (Veenman, 1995). This is in stark contrast to strictly mandated "multigrade" classrooms described by Mason and Burns (1996) where teachers must deliver "two different curricula to students of twice the age range" (p.313). These classrooms are "embedded in a graded system" where teachers without mixed-age training utilize single-grade teaching materials to prepare at least two grade-levels of students to go back into single-grade classrooms the following year (p.313). Understanding the circumstances that lead to mixed-age grouping provides insight into the level of systemic support behind the structure, which may affect availability of teacher training, teaching materials, and buy-in of both parents and educators, all critical components in an education paradigm (McEwan, 2008; Miller, 1991; Mulryan-Kyne, 2007). Circumstances that give rise to mixed-age grouping, then, have implications on its measured success.

Outcomes

One way to assess the effectiveness of a mixed-age program is by observing student outcomes. Numerous studies establish single-grade comparison groups in order to assess various effects of mixed-age grouping on students, but before presenting these findings, it is critical to problematize the use of student

outcomes to determine education practice. Firstly, two distinct groupings must be compared which, in the case of single-grade versus mixed-age grouping, implies numerous confounding variables (e.g., different teachers and classrooms). Rural students, for example, historically under-perform on outcome measures (Young, 1998). Secondly, what we use to measure outcomes is often related to curriculum content, as demonstrated by the use of standardized arithmetic and reading tests in the following studies. While these measures indicate the effect of our current education system and practices on student performance, they cannot offer insight into alternative practices that might, in fact, better promote accessible quality education. And finally, any conclusion of “better-ness” in education research implies “better for education” which is a weighty claim deserving of much justification. Why and how does performance on this particular measure indicate a quality education? Is the outcome measure reliable and valid across time and place? What behaviours and values more broadly are promoted and reflected by these research and education practices? These considerations will enable a more discerning illustration of the implications of presented studies.

Veenman (1995) consolidated evidence from over fifty studies on the cognitive and non-cognitive effects of “multigrade” classrooms – formed for “administrative and financial reasons.” It is not always clear how Veenman determined these studies were “multigrade” (Mason & Burns, 1996), but it seems to include reviews of rural schools and classes that mixed age groups in order to meet class size requirements. These classrooms are primarily default mixed-age or mandated to be mixed-age, and it is unclear how counterfactuals were established (e.g., whether rural single grade classrooms were compared against rural mixed-age classrooms). Cognitive outcomes were measured primarily with standardized tests of maths and reading. Findings on both cognitive and non-cognitive outcomes were mixed and effect sizes were small. The majority of studies identified no significant differences in mixed-age and single-grade student performance and about an equal number found either slight positive or negative significant results.

Veenman (1995) also examined eighteen studies on the cognitive and non-cognitive effects of “multiage” classrooms – formed for “pedagogical and didactic motives”. These studies all focused on classrooms where mixed-age was the preferred grouping method, though this does not clarify the manner in which mixed-age pedagogy was implemented. Veenman describes one school where “the principal wanted to move to complete multi-age grouping in order to foster more individualized and better instruction in the future,” but “the structure of the classroom and methods of instruction were left entirely up to the individual teachers” (Veenman, 1995, p.357). Other studies offered little detail on how or why mixed-age grouping was preferred, but there was presumably variability in the adoption of mixed-age methods across observed schools. Again, cognitive outcomes were measured primarily with standardized tests of maths and reading. In general, differences in cognitive outcomes did not reach significance. Non-cognitive outcomes indicated a slight positive effect of mixed-age grouping on “self-concept” and “attitudes toward school”, though six of the seven studies indicated these results lacked “evidence of initial equality” between comparison groups (p.366).

Berry (2001) explored reading improvement over a three-year period in low, medium, and high achieving third- through fifth-graders in Turks and Caicos. The mixed-age schools were from the least populated areas of the country, formed by default due to low numbers of students and teachers. Mixed-age schools had lower teacher-to-student ratios as well as teachers with less training and these factors were not controlled in the quantitative analysis. Berry utilized the McLeod reading test, which is not developed for specific grade levels. Low-achievers in mixed-age schools had significantly improved

reading scores compared to their single-grade counterparts. One subset of older, high-achieving students in mixed-age schools showed significantly less reading improvement than their single-grade counterparts.

The *Escuela Nueva* mixed-age reform program in Colombia has persisted for over forty years and has been imitated by other rural Latin American schools (McEwan, 2008). Geography and demographics render mixed-age the default grouping, but *Escuela Nueva*'s longevity, governmental support, and adoption of mixed-age practices outlined by Anderson and Pavan (1993) (e.g., flexible promotion) indicate intentional systemic action to support mixed-age grouping. McEwan (2008) reviewed the history and findings of mixed-age reforms in Colombia, Chile, and Guatemala. Results from Colombia's *Escuela Nueva* were the most robust and the only ones with matched control schools. Data indicated higher Spanish and maths scores for mixed-age third-graders compared to single-grade students in similar geographic areas and across all rural schools in the Pacific region of Colombia. This effect was not found in fifth-graders. Mixed-age students exhibited significantly higher levels of civic behaviour while there was no observed effect on creativity or self-esteem. Additionally, the successful "implementation of mixed-age reform was positively associated with a measure of peaceful interactions among students, even after controlling for other student and school variables" (p.477).

Gutierrez and Slavin (1992) conducted a comprehensive review of "nongraded" American schools, which examined both classrooms that mixed-age groups for preference and programs that resulted in age heterogeneity as a by-product of other grouping methods. Included studies utilized standardized tests to measure academic performance. A summary of findings on "Joplin-like nongraded programs" that created age heterogeneity by dividing students into reading ability groups across grade-levels revealed "substantial positive results in favour of the nongraded program" (p.348). Positive effects were amplified amongst programs that employed "flexibility in pupil grouping, with frequent assessment of mastery at each level, and increased amounts of teaching time for the homogeneous instructional groups" (p.348). Gutierrez and Slavin (1992) also evaluated findings on comprehensive nongraded programs that "emphasize continuous progress and flexible, multiage grouping" across all academic subjects (p.343). The majority of studies reviewed found significant positive results in favour of the mixed-age structure while none favoured the single-grade structure. The researchers conducted a separate evaluation of comprehensive nongraded programs that emphasized individualized instruction and individually guided education. Findings from these programs indicated no significant differences in student performance from single-grade structures.

Nye (1995) utilized standardized tests to measure academic performance in "nongraded" schools in Tennessee. She gives a lengthy explanation of what comprises a "nongraded" program, highlighting the importance of flexible grouping, but later states that the observed schools "had been validated as having the necessary components (to one degree or another) to generally be considered as nongraded" (p.10). This leaves considerable ambiguity in relation to particular techniques employed in the classrooms. Results showed that "nongraded" second-grade students significantly outperformed their single-grade counterparts in vocabulary, reading, language, and maths. First- and third-grade students scored significantly higher in vocabulary. Other results for first, second, and fourth grade students did not reach significance, but tended to favour the mixed-age structure. "Nongraded" third- and fourth-grade students scored significantly higher on a writing assessment test.

Two large-scale studies report outcomes of mixed-age groupings without reporting circumstance of grouping formation due to sheer number of schools and classrooms included in the investigation. Quail and Smyth (2014) conducted an extensive comparison of over 8,000 mixed-age and single-grade 9-year-olds from over 800 schools in Ireland. The researchers specify that urban, rural, and mixed schools were included and that they examined effects on students of being grouped with younger, older, or younger and older students. Using data collected from a national longitudinal survey, they were able to control for a variety of factors, including class size and teacher experience. Test results revealed no significant differences in reading or maths performance between mixed-age and single-grade classes, with and without controlling for extraneous variables. Investigation into gender effects revealed that girls tended to have more negative effects of mixed-age grouping than boys, exhibiting significantly lower reading and maths scores. Students in mixed-age classes self-reported significantly lower perceptions of their own behaviour, intellectual abilities, and popularity, and again, this effect was more pronounced amongst girls than boys. These negative perceptions were most extreme amongst girls grouped with older students.

Lindstrom and Lindahl (2011) utilized a national data set with information on more than 8,000 individuals to examine mixed-age versus single-grade classes in Sweden. Sixth-grade students, the eldest in their mixed-age classes, scored significantly lower on cognitive tests than their single-grade counterparts. By ninth-grade, however, there was no difference between mixed-age and single-grade students. Importantly, sixth-grade measures were results from a battery of tests, while ninth-grade results were averages of grades across different subjects in school which might reflect “students’ diligence and behaviour” (p.132).

Perspectives

Smit et al. (2015) observed that ambiguous empirical evidence “can be interpreted as either a positive or negative, depending on the expectations or fears that are connected with multigrading” (p.99). The unclear evidence in relation to mixed-age outcomes leaves the decision to mix age groups particularly susceptible to influences of attitude. Perhaps surprisingly, a history of slight and somewhat neutral effects on student outcomes does not seem to be accompanied by neutral attitudes. Generally, it seems parents view mixed-age as inferior to single-grade grouping. They worry that young students “won’t be able to keep up” and older students’ “advanced needs won’t be met” (Cornish, 2015, p.1). Wilkinson and Hamilton (2003) assessed various aspects of mixed-age grouping in New Zealand, where the majority of classrooms are mixed-age. Every mixed-age teacher reported that the most difficult aspect of their teaching was parent concerns about mixed-age grouping. Teachers also tend to hold negative attitudes toward mixed-age grouping, which may be explained by the widely acknowledged conviction that mixed-age grouping increases organizational and time demands on teachers (Mulryan-Kyne, 2007; Quail & Smyth, 2014; Seban, 2015; Smit & Engeli, 2015; Vincent, 1999). Teachers share parents’ fears about accommodating the learning of a diverse age range of students and worry that aspects of basic curriculum might get lost without the use of traditional teaching materials (Vincent, 1999). Bailey et al. (2016) identified general dissatisfaction with a mixed-age classroom until the implementation of a mixed-age program, including various teaching methods and trainings. Parents of students who attended mixed-age classes for multiple years reported the most positive attitudes toward mixed-age grouping. Parents also reported that their children enjoyed school more and performed well in this mixed-age grouping. Researchers’ conjectures and questions about mixed-age grouping reveal strong attitudes held amongst the academic community as

well. Huf and Raggl (2015) caution that mixed-age grouping is idealized by some after observing student interactions in their meta-ethnographic investigation of mixed-age classes. Hargreaves (2001) describes a lack of faith in multigrade pedagogy due to a pervasive belief that it is easier to teach students of the same age.

Discussion

The empirical evidence comparing mixed-age and single-grade outcomes is a veritable mixed bag, leading researchers to an array of contradictory conclusions. Veenman (1995) posits that mixed-age is “simply no better and no worse” than single-grade structures (p.367). Mason and Burns (1996) refute this conclusion pointing out that many educators purposefully funnel high-performing students and experienced teachers into mixed-aged classrooms to “ameliorate the difficulties and potential detriments involved” (p.312). They argue that, after accounting for this selection bias, neutral findings actually indicate at least a small negative effect. Other researchers assert that neutral findings suggest that the impact of other variables such as teaching quality transcend the impact of grouping structure (Berry, 2001; Quail & Smyth, 2014; Smit & Engeli, 2015; Veenman, 1995). Some practitioners, though critical, come down in support of mixed-age grouping (Stuart et al., 2007; Vincent, 1999), citing theories around collaborative learning from Vygotsky, Piaget, and Dewey as lending philosophical support for this approach.

Keeping in mind the circumstances that give rise to mixed-age grouping, it appears to me that a trend emerges. Examinations of intentionally created mixed-age classrooms, born of preference and with access to systemic support tend exhibit more favourable outcomes. Reforms in Colombia and Tennessee and comprehensive “nongraded” programs that were accompanied with specific mixed-age training and materials indicated positive impacts of mixed-age grouping (Gutierrez & Slavin, 1992; McEwan, 2008; Nye, 1999). This is quite a reasonable notion: groups with enthusiastic and well-resourced educators behind them have more positive effects on students. The extra effort required to elicit positive outcomes from mixed-age grouping illustrates a repeatedly cited (Aksoy, 2008; Hargreaves, 2001; Huf & Raggl, 2017; McEwan, 2008) point made by Little (2001) that “monograde organization of schools remains the taken-for-granted assumption of most of those who research and advise on curriculum development in both developed and developing countries” (p.491). She contends that “the knowledge, orientation and attitude required for effective multigrade teaching are invisible” inside today’s educational systems (p.490).

This widespread institutionalization of single-gradedness affects the validity of comparisons to mixed-age structures. One demonstration of this is the fact that the majority of researchers examining mixed-age outcomes utilized standardized tests created for specific grade levels to compare student performance (Gutierrez & Slavin, 1992; Lindstrom & Lindahl, 2011; Nye, 1995; Quail & Smyth, 2014; Veenman, 1995). With this in mind, I have to disagree with Mason and Burns’s (1996) argument that the overall neutrality of empirical results indicates negative effects. Rather, this suggests the untapped potential of mixed-age grouping as a viable option for today’s schools. Students in mixed-age classrooms generally perform as well on tests made specifically for their counterparts’ grouping structure, and in fact, the cause of the selection bias to which Mason and Burns (1996) refer is an anticipation of mixed-age “difficulties” and “detriments” on the part of administrators, revealing the underlying bias against mixing ages. Mixed-age grouping can hold its own in comparison

to its highly institutionalized rival structure, which leads me to wonder if systemic acceptance and support of this grouping would have currently underestimated or entirely unforeseen impacts.

Many researchers contend that the impacts of grouping structure in itself are superseded by other factors, including quality of teaching. Teaching quality is inarguably paramount in the provision of quality education, but teachers and students are inextricably linked and perpetually influenced by one another within the educational system. Gutierrez and Slavin's (1992) analysis indicates that individualized instruction may in fact neutralize the positive effects of a mixed-age program, supporting the notion that teaching approach and quality is indeed a critical factor in the success of mixed-age grouping. This finding also illustrates the necessity of implementing teaching methods that are specific to the grouping method. In order to understand the maximum potential of mixed-age grouping, we need to understand the teaching methods that provide the highest quality of education to mixed-age students. The importance of this need is reflected in the focus on mixed-age teaching strategies in recent mixed-age literature (Mulryan-Kyne, 2004, 2007; Roberts & Eady, 2012; Smit & Engeli, 2015). Lindstrom and Lindahl (2011) note that "there is little consensus about what characterizes this [the mixed-age] teaching method" (p.123). Commonly cited mixed-age practices include looping (i.e., assigning students to the same teacher for at least two years), differentiated instruction, flexible promotion, flexible within-class grouping, time-tabling, and teacher collaboration (Anderson & Pavan, 1993; Bailey et al., 2016; Gutierrez & Slavin, 1992; Hargreaves, 2001; Hoffman, 2002; Little, 2001; Smit & Engeli, 2015). Smit and Engeli (2015) seek to synthesize these practices in their empirical model of mixed-age teaching based on observations from small rural schools in Austria and Switzerland. Mulryan-Kyne (2007) discusses the importance of implementing mixed-age classrooms in conjunction with effective teaching methods and argues that, rather than implementing supplemental or separate mixed-age training, all educators should be trained in methods that account for high levels of diversity and heterogeneity amongst students. Bailey et al. (2016) offer a unique perspective on the importance of modified teaching methods in the successful implementation of mixed-age grouping. The researchers examined two American schools that initially established mixed-age classrooms in response to mandates around class-size requirements.

"The multigrade approach was disliked by staff and students due to its design which involved keeping the two grade levels separate within the classroom. This resulted in teachers who had to instruct one grade level while the other grade level worked independently...In addition, there was a deep concern for the amount of time wasted at the beginning of each year when a teacher had to spend much of the first month instructing the students in general classroom procedures as well as getting to know the individual learning capabilities of each student" (p.240).

In response to these difficulties, administration piloted a research-based mixed-age program that included teaching methods such as looping, differentiated instruction, and teacher collaboration. The school transitioned "all kindergarten through fifth grade classes to a multiage design" the following year (p.241). According to Stuart et al. (2007), teachers who implemented a mixed-age program in their classroom contend that they were able to develop this approach only because they had the freedom to choose a grouping method that aligned with their values and abilities. The successful implementation of mixed-age grouping requires much more than simply assigning a group of mixed-age students into a classroom that is structured as single-grade. It requires an alternative approach to teaching, and in order to realize its full potential, may require a new perspective on many aspects of education (e.g., school infrastructure, curriculum content, etc.). To summarize my review of mixed-age literature, I consider the most basic decision: to mix or not to mix.

To mix

Some primary schools mix age groups for practical reasons. When small enough numbers of students or teachers require it, schools will mix ages to create full classes. Other schools may mix age groups to maintain small and even class sizes. Some empirical evidence suggests positive outcomes, especially regarding social measures (Gutierrez & Slavin, 1992; Nye, 1995). Mixed-age grouping has been presented as a cost-saving structure due to its inherent flexibility, particularly in developing nations (Benveniste & McEwan, 2008; Berry, 2001; Lindstrom & Lindahl, 2011; Little, 2001). Theoretical arguments for mixed-age grouping abound. Huf and Raggl (2015) posit that in today's schools "the variety of social, familiar, cultural, and ethnic experiences children bring to school is huge...Against the background of a growing diversity of childhood experiences, age-mixed groups are rather considered as a catalyst for acknowledging this diversity" (p.232). Many other researchers share this notion that mixed-age grouping necessitates recognition of individual student identity (Hargreaves, 2001; Smit & Engeli, 2015). It is also suggested that mixed-age structures allow for collaborative and peer learning, crucial themes in the philosophies of both Vygotsky and Dewey (Benveniste & McEwan, 2008; Roberts & Eady, 2012). There is a lack of scholarly literature supporting the single-grade structure, and its implicit assumptions that age dictates how and what children should be learning is problematic (Anderson & Pavan, 1993; Vincent, 1999). Mixed-age groups are more reflective of the social groups that children will encounter throughout their lives.

Not to mix

Since the turn of the 20th century, single-gradedness has become the de facto standard of primary education (Aksoy, 2008; Little, 2001; Miller, 1993; Vincent, 1999). Its implementation surged with the rise of industrialization and accompanying increases in demand for equitable education access and well-trained teachers (Miller, 1993). Grouping students by age divides students along inarguable "pre-set, asymmetric roles" (Huf & Raggl, 2015, p.237), making education accessible to the masses while side-stepping debate on how to group students. Implementation of effective mixed-age grouping is expensive, demanding new resources and increased human labour (Anderson & Pavan, 1993; Little, 2001). There is a general consensus amongst researchers that other factors, like teaching quality, have greater impact on learning than grouping structure. Mixed-age teaching puts increased time-demands on teachers, necessitating more planning and organization. Implementing mixed-age structure requires different teacher training, teaching materials, and evaluation methods than those that are widely used today. The combination of slightly positive or negative but mostly neutral empirical findings and the apparent low impact of grouping method in itself simply does not justify the time and resources needed to replace single-graded education.

Future Directions

Whether mixed-age grouping is an unfavourable inevitability or an ethical obligation, circumstances ensure that it is here to stay in some capacity. In fact, many researchers predict that the prevalence of mixed-age grouping is likely to increase in the future (Huf & Raggl, 2017; Little, 2001; Mulryan-Kyne, 2007). Smit et al. (2015) reflect on the current condition of mixed-age research, stating that:

“research in the area of mixed-age education is still developing. One challenge is the present ambiguity in definitions of multi-grade education. Therefore, before research outcomes can reliably indicate the specific aspects of learning in mixed-age environments that are most beneficial, careful attention must be given to the definition and selection of multi-grade classrooms and detailed descriptions of teaching strategies should be provided” (p.102).

It is critical to observe and report the circumstances and methods of mixed-age classrooms in scholarly research. This will enable further examination and comparison to determine promising aspects of mixed-age grouping. It is also necessary to consider how to reliably and validly measure mixed-age phenomenon without relying on measures constructed within and for a single-grade system. Wilkinson and Hamilton (2003) utilized an ungraded reading test to examine the range of abilities in mixed-age and single-grade classes in New Zealand. They found no significant differences in the ranges of abilities within mixed-age and single-grade classrooms, perhaps supporting the idea that the assumed homogeneity in single-grade classrooms is an illusion that mixed-age grouping might work to dispel. The potential of mixed-age classrooms to make diversity evident is often cited as a reason to mix ages, but does this grouping actually have this effect? Does teacher acknowledgement of classroom diversity influence teaching strategies and subsequent student outcomes? Continued exploration of mixed-age teaching methods will help clarify how to maximize benefit of mixed-age grouping. Gutierrez & Slavin’s (1992) analysis indicated that individualized education tampered positive effects of the mixed-age program, while Smit & Engeli’s (2015) mixed-age teaching model is built upon the practice of differentiated education. How will teachers balance differentiated education with opportunities for collaborative learning? Mixed-age teachers face this and other dilemmas within the relatively uncharted territory of mixed-age grouping. So, if mixed-age classrooms are mandated, how will teachers be assigned to these potentially fraught mixed-age classrooms? Mason and Burns (1996) note that administrators often assign more experienced teachers to mixed-age classrooms, but Benveniste and McEwan (2000) observed that experienced teachers were less likely to adopt and practice new pedagogical techniques, which appear critical for successful mixed-age implementation.

Educators need not view mixed-age grouping as only classroom-level. Age heterogeneity can be instituted in primary schools at any time by simply enabling students to work in mixed-age groups for any duration, and indeed, flexible grouping practices are associated with positive outcomes in mixed-age research. A recent study by Kallery and Loupidou (2016) found that younger students’ science learning improved as a result of weekly mixed-age, small group study sessions. Future research could assess effects of isolated mixed-age interventions during the school day. Some mixed-age literature explores the perspectives of parents and teachers, but the voices of students are noticeably absent. Indirect measures of academic and social factors (e.g., test scores, ratings of self-efficacy, popularity, etc.) offer limited insight children’s experience inside a mixed-age classroom. Future inquiry could offer crucial understandings from the student perspective. Do they feel safe in mixed-age groups? Researchers and practitioners consistently report that effective collaboration across age groups takes time (Kallery & Loupidou, 2016; Roberts & Eady, 2012). Huf and Raggl (2015) document interactions between older and younger students in a mixed-age classroom, noting that older students seem to exercise power by giving directives and critical feedback. Does age heterogeneity reinforce hierarchies absent in a single-grade class or are these power differentials typical inside primary classrooms more generally? How are students’ ideas about education and learning altered by mixed-age grouping? Berry and Little (2007) suggest looking into more longitudinal outcomes as a means of evaluation, such as how mixed-age students perform at subsequent schools. Retrospective views from former mixed-age students could

offer an interesting perspective on how mixed-age grouping might shape lives, relationships, careers, etc. Future research should aim to paint a more holistic picture of the formation and operation of mixed-age classrooms and their positioning within broader societal and geographic systems. The decision whether to mix ages in primary schools is contextualized and complex and affects both accessibility and quality of education. This decision is deserving of the utmost criticality and intentionality.

Conclusion

Straying from the traditional graded model raises questions not only about how to define and enact quality teaching, but also about how we as a society define and enact “quality education.” At the heart of the scepticism toward mixed-age grouping is the fear of not measuring up to the expectations that a single-grade society has set as valid indicators of a good education (e.g., grades, standardized test results, certain social measures). Indeed, Lindstrom and Lindahl (2011) found that mixed-age teachers were significantly less likely to emphasize homework, basic knowledge, and formal testing. They were also significantly more likely to believe in students’ influence over their own education. Mixed-age students were found to have more negative views of their own behaviour and believe themselves to be less popular (Quail & Smyth, 2014). Does this suggest inferiority of mixed-age grouping? Does a “quality education” lead students to believe that they are popular? Altering education practice feels high stakes. It ignites the concerns of teachers, parents, researchers, politicians. What will happen? What will change? And how will we know? Even in the case of mixed-age grouping, whose outcome effects are shown to be quite comparable, people bristle at the idea of education reformation. And yet, current practice is not necessarily the best practice for transferring quality education, and in order to see other options, we have to look out for them, measure them, establish some level of confidence that they might add value. Mixed-age literature hints at numerous uncommon education practices that appear promising. I argue in favour of further inquiry and considered application. When it comes down to it, we must organize the population in such a way to allow their education, and to ensure quality, we must continuously reflect upon society’s goals and conduct critical research to incorporate these goals into practice.

References

- Aksoy, N. (2008). Multigrade schooling in Turkey: an overview. In *International Journal of Educational Development*, 28(2), 218–228. Retrieved from: doi: 10.1016/j.ijedudev.2007.05.002
- Anderson, R. H., & Pavan, B. N. (1993). *Nongradedness: Helping it to happen*. Lancaster, PA: Technomic Publishing Co.
- Bailey, G. J., Werth, E. P., Allen, D. M., & Sutherland, L. L. (2016). The Prairie Valley Project: reactions to a transition to a schoolwide, multiage elementary classroom design. In *School Community Journal*, 26(1), 239–264.
- Benveniste, L. A., & Mcewan, P. J. (2000). *Constraints to implementing educational innovations: the case of multigrade schools*, 18.

- Berry, C. (2001). Achievement effects of multigrade and monograde primary schools in the Turks and Caicos Islands. In *International Journal of Educational Development*, 21(6), 537–552. Retrieved from: doi: 10.1016/S0738-0593(01)00014-1
- Berry, C., & Little, A. W. (2007). *Multigrade teaching in London, England. Education for all and multigrade teaching: Challenges and opportunities* (pp. 67–86). Dordrecht, Netherlands: Springer.
- Bradbury, A. (2018). The impact of the Phonics Screening Check on grouping by ability: a ‘necessary evil’ amid the policy storm. In *British Educational Research Journal*, 44(4), 539–556. Retrieved from: doi: 10.1002/berj.3449
- Cornish, Linley. (2015, April 6). *Are mixed-grade classes any better or worse for learning*. Retrieved from: <https://theconversation.com/are-mixed-grade-classes-any-better-or-worse-for-learning-38856>.
- Di Lorenzo, L. T., & Salter, R. (1965). Co-Operative Research on the Nongraded Primary. In *The Elementary School Journal*, 65(5), 269–277.
- The Global Education 2030 Agenda*. (2015, September). Retrieved from: <http://www.unesco.org/new/en/santiago/education-2030/>.
- Gutiérrez, R., & Slavin, R. E. (1992). Achievement effects of the nongraded elementary school: a best evidence synthesis. In *Review of Educational Research*, 62(4), 333–376. Retrieved from: doi: 10.3102/00346543062004333
- Hargreaves, E. (2001). Assessment for learning in the multigrade classroom. In *International Journal of Educational Development*, 21(6), 553–560. doi: 10.1016/S0738-0593(01)00015-3.
- Huf, C., & Raggl, A. (2015). Social orders and interactions among children in age-mixed classes in primary schools – new perspectives from a synthesis of ethnographic data. In *Ethnography and Education*, 10(2), 230–241. Retrieved from: doi: 10.1080/17457823.2015.1017514
- Huf, C., & Raggl, A. (2017). The normativity of the helping child – meta-ethnographic perspectives on individualised learning in age-mixed classrooms. In *Ethnography and Education*, 12(2), 165–177. Retrieved from: doi: 10.1080/17457823.2016.1192479
- Hoffman, J. (2002). *Flexible Grouping Strategies in the Multiage Classroom*, 41(1). Retrieved from: <http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=11&sid=3b98e8cd-e119-45c5-ab0d-b4d87aec5f74%40pdc-v-sessmgr05>.
- Kallery, M., & Loupidou, T. (2016). Learning science in small multi-age groups: the role of age composition. In *International Journal of Science Education*, 38(9), 1570–1590. Retrieved from: doi: 10.1080/09500693.2016.1201871.
- Kelly-Vance, L., Caster, A., & Ruane, A. (2000). Nongraded versus graded elementary schools: an analysis of achievement and social skills. In *Alberta Journal of Educational Research; Edmonton*, 46(4), 373.
- Lindström, E., & Lindahl, E. (2011). The effect of mixed-age classes in Sweden. In *Scandinavian Journal of Educational Research*, 55(2), 121–144. Retrieved from: doi: 10.1080/00313831.2011.554692
- Little, A. W. (2001). Multigrade teaching: towards an international research and policy agenda. In *International Journal of Educational Development*, 21(6), 481–497. Retrieved from: doi: 10.1016/S0738-0593(01)00011-6
- Mason, D. A., & Burns, R. B. (1996). “Simply no worse and simply no better” may simply be wrong: a critique of Veenman’s conclusion about multigrade classes. In *Review of Educational Research; Washington*, 66(3), 307.
- McEwan, P. J. (2008). Evaluating multigrade school reform in Latin America. In *Comparative Education*, 44(4), 465–483. Retrieved from: doi: 10.1080/03050060802481504
- McGillicuddy, D., & Devine, D. (2018). “Turned off” or “ready to fly” – ability grouping as an act of symbolic violence in primary school. In *Teaching and Teacher Education*, 70, 88–99. Retrieved from: doi: 10.1016/j.tate.2017.11.008
- Miller, B. A. (1991). A review of the qualitative research on multigrade instruction. In *Journal of Research in*

Rural Education, 7(2), 3 – 12.

Miller, B. A. (1993). A review of the quantitative research on multigrade instruction. Multiage classrooms: The ungrading of America's schools. In *The multiage resource book* (pp. 65 – 83). Peterborough, NH: Society for Developmental Education.

Mulryan-Kyne, C. (2004). Teaching and learning in multigrade classrooms: what teachers say. In *The Irish Journal of Education*, 35, 5 – 19.

Mulryan-Kyne, C. (2007). The preparation of teachers for multigrade teaching. In *Teaching and Teacher Education*, 23, 501 – 514.

Nye, B. A., & And Others. (1995). *Are multiage/nongraded programs providing students with a quality education? Some answers from the school success study*. Paper presented at The Fourth Annual National Create the Quality Schools Conference. Retrieved from: <https://eric.ed.gov/?id=ED384998>.

Quail, A., & Smyth, E. (2014). Multigrade teaching and age composition of the class: the influence on academic and social outcomes among students. In *Teaching and Teacher Education*, 43, 80–90. Retrieved from: doi: 10.1016/j.tate.2014.06.004

Roberts, J., & Eady, S. (2012). Enhancing the quality of learning: what are the benefits of a mixed age, collaborative approach to creative narrative writing? In *Education 3-13*, 40(2), 205–216. Retrieved from: doi: 10.1080/03004279.2010.511624

Seban, D. (2015). Development of preservice identities: learning from a multigrade classroom practicum context. In *Journal of Education for Teaching*, 41(1), 19–36. Retrieved from: doi: 10.1080/02607476.2014.992632

Slavin, R. E. (1988). *Synthesis of Research on Grouping in Elementary and Secondary Schools*, 13.

Smit, R., & Engeli, E. (2015). An empirical model of mixed-age teaching. In *International Journal of Educational Research*, 74, 136–145. Retrieved from: doi: 10.1016/j.ijer.2015.05.004

Smit, R., Hyry-Beihammer, E. K., & Raggl, A. (2015). Teaching and learning in small, rural schools in four European countries: introduction and synthesis of mixed-/multi-age approaches. In *International Journal of Educational Research*, 74, 97–103. Retrieved from: doi: 10.1016/j.ijer.2015.04.007

Stuart, S. K., Connor, M., Cady, K., & Zweifel, A. (2007). Multiage instruction and inclusion: a collaborative approach. In *International Journal of Whole Schooling*, 3(1), 12–26.

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: a best-evidence synthesis. In *Review of Educational Research*, 65(4), 319–381. Retrieved from: doi: 10.3102/00346543065004319

Vincent, S. (Ed). (1999). *The multigrade classroom: a resource handbook for small, rural schools*. Portland, OR: Northwest Regional Educational Library.

Wilkinson, I. A., & Hamilton, R. J. (2003). Learning to read in composite (multigrade) classes in New Zealand: teachers make the difference. In *Teaching and Teacher Education*, 19(2), 221–235. Retrieved from: doi: 10.1016/S0742-051X(02)00105-1

Young, D. J. (1998). Rural and Urban Differences in Student Achievement in Science and Mathematics: A Multilevel Analysis. In *School Effectiveness and School Improvement*, 9(4), 386–418. <https://doi.org/10.1080/0924345980090403>.