Alberta Journal of Educational Research, Vol. 68.2, Summer 2022, 214-233

The Practice of Excluding Students from Large-scale Assessments: Interviews with Principals

Tess Miller[1], Anne Marie FitzGerald[1], Han Xu[2]

1 University of Prince Edward Island, 2 Queen's University

Given the shortage of literature examining exclusion rate practices in Canada, this study reported on exclusion rate practices from principals' perspectives. Specifically, factors influencing principals' decision-making regarding excluding students from large-scale assessments (LSAs) were explored. We inquired about factors and implications of social promotion as well as decision-making based on compassion, particularly in relation to students who were working below grade level. Using a qualitative research design, we conducted interviews with eight school principals and one vice-principal. Findings revealed an absence of clarity and governance in LSAs that likely influenced differences in principals' decision-making. Some principals would automatically exclude students who had an Individualized Education Plan, while others would review student files to determine which students to exclude. There was also a discrepancy surrounding which educational supports were permitted on the LSAs.

Étant donné le manque de documentation sur les pratiques en matière de taux d'exclusion au Canada, cette étude porte sur les pratiques en matière de taux d'exclusion du point de vue des directeurs d'école. Plus précisément, nous avons exploré les facteurs qui influencent la prise de décision des directeurs d'école concernant l'exclusion des élèves des évaluations à grande échelle (EGE). Nous nous sommes renseignées sur les facteurs et les implications de la promotion sociale ainsi que sur la prise de décision fondée sur la compassion, en particulier en ce qui concerne les élèves qui travaillent en dessous du niveau scolaire. À l'aide d'un modèle de recherche qualitative, nous avons mené des entretiens avec huit directeurs d'école et un directeur adjoint. Les résultats ont révélé un manque de clarté et de gouvernance dans les EGE, ce qui a probablement influencé les différences dans les décisions des directeurs d'école. Certains directeurs d'école excluaient automatiquement les élèves qui avaient un plan d'enseignement personnalisé, tandis que d'autres examinaient les dossiers des élèves pour déterminer lesquels exclure. Il y avait également une divergence quant aux soutiens éducatifs autorisés dans le contexte des EGE.

Principals perform complex work that involves cultivating a shared vision, leading instruction, creating a safe and inclusive climate, and ensuring accountability for student learning (Pollock et al., 2017). Accountability for student learning in Canada focuses on improving instruction and student achievement (Klinger et al., 2008; Volante & Ben Jaafar, 2008). In this context, Canadian school principals respond to multiple mandates regarding inclusive practices for students with special education needs in the administration of large-scale assessments (LSAs) at the international, national, and provincial levels (Pollock & Winton, 2016). Although units of analysis

differ within each level of LSA, each LSA purports to provide an indicator measuring the wellbeing of the educational system, direct educational resources to where they are needed most, inform educational policy, and ultimately, improve student learning. Often, LSA achievement results are anxiously awaited, widely publicized, and utilized politically to promote various means of school improvement (Grek, 2009).

Stakeholder confidence in LSA results depends on schools' adherence to administration guidelines, such as identifying which students must write the assessment and which students may be excluded. International and national bodies suggest a maximum 5% exclusion rate and provide guidance for how to make exclusion decisions (Brzyska, 2018). In provincial and territorial jurisdictions, LSAs and exclusion criteria are determined locally thus, exclusion practices may vary. High rates of exclusion in some provinces and territories contributed to Canada's overall exclusion rate of 6.9% on the Program for International School Assessment (PISA; CMEC, 2018). In this LSA and the Pan-Canadian Assessment Program (PCAP), students can be excluded from writing the assessment for one of the three reasons: (1) functional or physical disability that prevents the students from completing the assessment, (2) limited language skills that hinder students from reading or speaking the language of the assessments, and (3) intellectual disability such as a mental or emotional disability that has resulted in a cognitive delay hindering students' ability to complete the assessment (CMEC, 2018; OECD, 2017).

Given that there is a wide spectrum of intellectual disabilities, determining which intellectually disabled students are capable of writing the LSA and which students are not presents many challenges. These challenges are amplified given that nations around the world are moving towards more inclusive education practices for these students with special education needs and Canada is no exception (Brzyska, 2018). The Canadian Charter of Rights and Freedoms guarantees an inclusive education to children with exceptionalities; however, each province and territory determines how such education will be structured and delivered. An estimated 15% of Canadian school children are considered students with exceptionalities, encompassing differences in behaviour, communication, physical ability, and cognition (Timmons, 2008). Inclusive practice deems that all children be included to the extent possible. Efforts to include students with exceptionalities may result in the school's creation of an Individualized Education Plan (IEP), an Academic Learning Plan (ALP), or other modifications to the curriculum or adaptations for the child. Adaptations may include supports such as number lines, voice-to-text software, or writing on computers. Decisions to include or exclude children with exceptionalities in LSAs or to permit adaptations are made by principals, typically in conjunction with classroom teachers, special education teachers, other support personnel, district administrators, and parents, often on a case-by-case basis.

There are many reasons, other than exceptionality, to explain why students have been excluded from LSAs. For example, students can be excluded from an assessment and described as a *non-participant* for three reasons: (a) parental refusal to participate in the assessment, (b) students were no longer in the school, or (c) the age criterion was not matched (CMEC, 2018). Over the past four cycles of the PISA, there has been great variability in the range of non-participating students. On the 2018 PISA, the range of non-participating students spanned from 0–3.7%; on the 2015 PISA, the range was 2.8–35.7%; on the 2012 PISA, the range was 4.0–16.8%; and on the 2009 PISA, the range was 0–26%. These wide ranges of non-participating students could affect the reliability of the assessment as there is uncertainty as to whether the sample is a true representation of the population. PISA administrators have acknowledged concerns related to the representativeness of their samples, stating: "Further steps will be required in future PISA

cycles to address the issue of high exclusion rates for schools and students in some provinces" (CMEC, 2019, p.82).

On the national PCAP, students can also be excluded under a category titled *other*. In this category, students can be excluded:

> (1) who are exempted by the school; (2) who are exempted because appropriate modifications could not be made; (3) who are no longer enrolled in the selected school/class; (4) who do not wish to participate; and (5) for whom no data were available (CMEC, 2018, p.156).

Unfortunately, there is no additional detail describing what criteria the school can use to exempt a student or how many students choose not to write the assessment, assuming a choice was presented. These five subcategories presented under the *other* heading for excluding students on the national LSA, clouds the transparency of this LSA. On the 2016 PCAP, the percent of students excluded from the assessment under the category of *other* ranged from 1.6–6.3% (CMEC, 2018). In prior cycles of the PCAP, the percent of excluded students ranged from 2.1–12.5% in 2013 and 0.3–1.3% in 2010 (CMEC, 2011, 2014). On the 2013 PCAP, the percent of students excluded in the *other* category, was almost double the percent in 2017 and on the 2010 PCAP, the percent was a fraction of the percent in 2017. These fluctuations in the percent of students excluded falling into the *other* category warrant further investigation.

Exclusion criteria on provincial LSAs are similar to national and international exclusion criteria (Miller & Yan, 2022). However, in the provincial jurisdiction we examined, students could be exempted from the LSA if students were on one of eight different programs or learning plans: alternative education, English as an Additional Language (EAL), exemption (extenuating circumstances), IEP, modified program, learning plan, modified program, and transition action plan (Miller & Yan, 2022). Also, the number of non-provincial writers increased when the number of students who were absent on the day of the assessment or refused to participate in the assessment (their own decision or their parents' decision) were added to the total. From 2011 to 2018, the percent of students excluded from the provincial assessment ranged from 8.4%–13% (Miller & Yan, 2022).

This study builds on a previous study (Miller & Yan, 2022) that connected high rates of exclusion on LSAs to social promotion. The focus of this study was to understand factors influencing principals' decision-making when deciding to exclude a student from a LSA. In particular, we wanted to know about decisions related to excluding students from LSAs who had intellectual disabilities or fell into the non-participant and *other* categories as described above. Previous research has shown the practice of excluding students has been fraught with misunderstanding due to an absence of clear guidelines and consistency from jurisdiction to jurisdiction (Brzyska, 2018; Gamazo et al., 2019; LeRoy et al., 2019) as well as biasing to raise the perception that the school is doing well (Albrecht & Joles, 2003; Braun et al., 2010; Brzyska, 2018; Emler et al., 2019; Prais, 2003). It is suspected that with a growing connection between LSA results and school accountability for the achievement of all students, combined with principals' personal reasons for excluding students from LSAs as well as the absence of guidelines, transparency, and standards in excluding students, student samples might not reflect the true population, thus raising the overall perception of the school (Braun et al., 2010; Brzyska, 2018; Miller & Yan, 2022; Prais, 2003; Smith & Douglas, 2014). Hence, the primary research question focused on factors influencing the representativeness of LSA participation. Specifically, the research question posed in this study was: what factors influence principals' decision-making to

include or exclude students from LSAs.

This research is framed within the lens of accountability, where those responsible for educating children are held accountable for their learning as determined by achievement scores on LSAs (Leishman, 2015; Smith & Douglas, 2014). In the context of this study, accountability for student learning lies with school principals. Principals have adopted different leadership styles to respond to the challenges of LSAs (Prytula et al., 2013; Schulze, 2014). At the same time, some principals find it stressful to be held accountable for student learning (Prytula et al., 2013; von der Embse et al., 2016). Accountability for student learning is even more paramount in schools that practice social promotion, which has resulted in a group of students whose learning has fallen well below their age-appropriate peers (Miller & Yan, 2022). However, schools that combined pedagogical autonomy with increased accountability reported better student performance (Bédard, 2015; OECD, 2011). It is uncertain as to whether this sense of autonomy influenced principals' decisions to exclude students more liberally thus, creating a sampling bias (Brzyska, 2018). As argued by Klieger (2016), principals' understanding of LSA results would influence the assessment culture and success within schools. Therefore, there is much to learn by interviewing principals about their decision-making related to excluding students from LSAs.

We posit that principals exclude students from writing LSAs more liberally than permitted for reasons related to compassion and implications of social promotion. Social promotion is a school practice where students are passed to the next grade with their peers even though they failed to achieve the academic requirements or meet performance mandates (Doherty, 2004; McMahon, 2018). However, students who have been socially promoted to the next level, if not receiving proper accommodation, would face challenges in achieving the grade level requirements (Chataa & Nkengbeza, 2019). The Principles for Fair Student Assessment Practices for Education in Canada (PFAPEC) provided support to exclude students from writing LSAs if they were deemed to be operating too many grade levels below the level being assessed (Rogers, 1993). The PFAPEC is an outdated and contentious document given that some researchers have advocated that all students should write LSAs; otherwise, it is not known which schools and students are in need of assistance (Thompson et al., 2002; Zumeta, 2015). Conversely, it is problematic having students who have been promoted two or three grade levels, write a LSA that they do not have the knowledge and understanding to write. Including such students in a LSA may have a dire impact on their self-worth, which is probably at the forefront of students' minds already if they are so far behind their peers. In these cases, principals will likely exercise compassion and exclude these students from the LSA.

Literature Review

This literature review is organized into three sections. The first section focuses on general research about exclusion practices and is followed by a brief introduction to social promotion and research that connects social promotion to exclusion rates. Last, we synthesized research related to principals' role in LSAs and factors influencing their decisions to include or exclude students from LSAs.

Research on Exclusion Practices

Previous literature has identified several factors contributing to high exclusion rates of which many factors stemmed from vague exclusion guidelines and a shortfall in practices documenting

the extent to which exclusion guidelines have been implemented (Anderson et al., 2011; Braun et al., 2010; Eivers, 2010; LeRoy et al., 2019; McGrew et al., 1992; Miller & Yan, 2022; Murphy, 2014; Nichols et al., 2006; Prais, 2003; Thurlow, 1995; Wuttke, 2007; Ysseldyke & Thurlow, 1994). Disagreements over the definitions of disabilities, ways of processing cognitive versus academic weaknesses, as well as various diagnostic models for identifying learning disabilities further muddied the ambiguity surrounding who should be excluded from LSAs (D'Alessio & Watkins, 2009; Gamazo et al., 2019; LeRoy et al., 2019). The current identification criteria for students with special education needs within PISA vary across countries with heavy reliance on school administrators' judgment (Gamazo et al., 2019; LeRoy et al., 2019; Schuelka, 2013). Consequently, the incomparability of criteria creates methodological difficulty in the interpretation of the LSA results. Early on, Wuttke (2007) had discussed issues within the PISA regarding the vague exclusion criteria and varied exclusion rates among participating countries, thus putting the PISA statistics and its outcome under critique. As of the 2018 PISA report, issues with exclusion have persisted (CMEC, 2019).

Despite these issues, LSAs such as the PISA and PCAP have captured a growing interest from researchers, as they have been increasingly used to hold schools accountable for students' achievement and for engaging in educational reform (Leishman, 2015; Smith & Douglas, 2014). LSAs are often used as measures that hold schools accountable for progress in students' achievement (Smith & Douglas, 2014). In this environment, students who struggle with learning might be viewed as draining school resources and hindering schools' overall competitiveness (Graham & Jahnukainen, 2011). Empirical evidence in the Canadian context has also shown that students with disabilities performed worse than others in LSAs (Hinton, 2014). Further, it has been noted that the loosely defined exclusion criteria might create some leeway for participating schools to control the participation of students, thus causing an upward bias in reporting school achievement (Braun et al., 2010; Brzyska, 2018; Gamazo et al., 2019; Prais, 2003; Wuttke, 2007). Braun et al. (2010) challenged the credibility of LSA results based on the connection between the substantial variation in exclusion rates and students' performance, noting, the worse a student performed, the higher the probability the student was excluded. For example, in the 2000 PISA report, evidence was found that the underrepresentation of a particular group of Austrian students led to inflated scores (Eivers, 2010). Given what has been discussed so far, it is reasonable to infer those jurisdictions with high exclusions are likely to have inflated achievement scores given that achievement scores are not likely to represent the true population of students, which is contrary to the accurate representation of the achievement called for by Zumeta (2015). In a more recent study, Anders et al. (2021) provided more evidence to substantiate concerns related to exclusion rate practices. These researchers also challenged the creditability of LSA results in their examination of Canada's 2015 PISA scores and concluded that Canada's PISA data was plagued with "high student exclusion rates, low levels of school participation rates, and high rates of student absence" (p. 246). Available studies called for greater transparency in practices related to excluding students from writing LSAs as well as better guidelines governing this practice (Brzyska, 2018; Eivers, 2010; Gamazo et al., 2019).

Social Promotion

Social promotion is a school practice where students are passed from one grade to the next with their peers even though they failed to achieve the academic requirements or met performance mandates (Doherty, 2004; McMahon, 2018). Historically, decisions to promote or retain a

student in the unsuccessful grade were dependent on the philosophy of schools and school administrators (Ayers, 1909). Given the impact of social promotion or retention on students' wellbeing, this practice has been a subject of long-standing debate (Ahmed & Mihiretie, 2015; Anastasiou et al., 2017; Crepeau-Hobson, 2016; Hwang & Cappella, 2018; McMahon, 2018). Social promotion can work if the student receives remediation and an opportunity to learn in the following grade (Connor, 2018; Wanzek et al., 2020). However, the burden of receiving remediation on prior concepts combined with learning new concepts may thwart intentions to minimize impact on self-esteem (Ahmed & Mihiretie, 2015). Subsequently, students who have been unsuccessfully socially promoted fall below grade-level standards (Chataa & Nkengbeza, 2019). If these students participate in LSAs, it is likely that they will hold back school achievement, which presents tension for school principals who are ultimately accountable for achievement as measured by LSAs (Elliott & Thurlow, 1997; Smith & Douglas, 2014).

It is unknown to the academia how widely social promotion has been practiced in the Canadian provinces and territories. Despite a latent connection between exclusion practices and social promotion, there is an absence of literature focusing on LSA exclusion practices in the Canadian context combined with an absence of research that links social promotion practices to exclusion rates on LSAs (Miller & Yan, 2022).

Principals' Role in LSAs

Current literature involving school principals and LSAs have primarily focused on their role and leadership in implementing LSAs, while very few researchers underscored the importance of principals' perspectives and decision making related to excluding students, particularly those with special needs (Heydrich et al., 2013; Newton et al., 2010; Prytula et al., 2013; Renihan & Noonan, 2012; Schleicher, 2009; Sealy et al., 2016). From this perspective, principals are faced with a quandary between their role as instructional leaders and their role in providing leadership in the area of LSAs (Newton et al., 2010; Prytula et al., 2013; Renihan & Noonan, 2012; Sealy et al., 2016). For example, in Newton et al. (2010)'s study involving interviews with 25 school principals, they found that LSAs had reshaped the role of principals into supporting teachers, facilitating and understanding LSAs, as well as monitoring interventions. They also reported that the increasing focus on LSAs and accountability had placed significant pressure on principals, a finding that was corroborated in two other studies (Prytula et al., 2013; Renihan & Noonan, 2012). Prytula et al. (2013) argued that overall, LSAs had a positive influence on principals through transforming them into instructional leaders and improving teaching and learning. However, the decentralized feature of Canada's education system has led to varied accountability contexts from province to province, resulting in different levels of principals' engagement in facilitating LSAs (Renihan & Noonan, 2012). Further differences were reported by Renihan and Noonan (2012), who noted that contextual differences would influence principals' leadership and involvement in LSAs. For example, small schools might question the value of their data in LSAs since it tended to become less significant in comparison with that of larger schools, which might eventually have an impact on small schools' philosophy of participating in LSAs. Hence, we anticipate that our interviews with principals about their decisions to include or exclude students from LSAs will also vary from one school to the next, given the contextual differences between French and English school boards as well as the rural setting of schools which is all underpinned by vague exclusion guidelines.

In summary, there is an absence of literature on principals' perspectives on exclusion criteria and LSA practices embedded within the context of inclusion and social promotion. Given the

variability surrounding definitions of disabilities combined with the common practice of social promotion and high exclusion rates reported on LSAs, it is important to understand factors influencing principals' decision-making to address issues of high exclusion rates and the interplay of inclusion and social promotion. Findings from this study have the potential to impact policy related to exclusion criteria.

Methodology

In this qualitative study, eight public school principals plus one vice-principal from one Atlantic Canadian province were selected to participate in a 1-hour interview. In this provincial jurisdiction, the purpose of the provincial LSA in grades 3, 6, 9, and 10 is for accountability, gatekeeping, and monitoring of student achievement (Klinger et al., 2008). More specifically, the provincial grade 3 and 6 assessments are designed to inform teachers, parents, and students as well as improve teaching and learning and guide professional development. These purposes of the provincial LSAs apply to the grade 9, 10, and 11 LSAs and in addition, the grade 9 and 10 LSAs can be viewed as gatekeepers given that students' scores on the grade 9 assessment of mathematics contribute 10% of their final grade; however, the LSA score does not prevent a student from achieving a passing score of 50% in the course (Miller, 2013). In contrast, the grade 10 literacy assessment was a graduation component, and if students were unsuccessful, they have an opportunity to receive intervention and re-write the assessment. The grade 11 LSA of mathematics can be considered a higher-stake gatekeeper in that students' scores contributed 25% of their final grade in a course (Yan, 2016).

We applied a stratified random sampling technique where a minimum of two schools were selected to represent students in grade 3 and 6 (primary/elementary), grade 9 (intermediate), and 10 (senior) to ensure opportunities for principals to discuss exclusion practices on provincial, national, and international assessments. Subsequently, principals were from a range of school divisions who participated in at least one LSA. One principal was from a French-language school and another principal was from a school with a French immersion program. The enrollment in schools ranged from approximately 200 (K–6 school) to 600 (10–12 high school) for an average school population of 386 students (Anonymous, 2020). Five participants were female and participants' experiences as principals ranged from one year to 15 years. Interviews were conducted over two months in February and March 2020.

Interviews were semi-structured and ranged from approximately 40 to 60 minutes. One interview was conducted using the software Zoom and the others were in-person at the participants' school. Interview questions were scenario-based and open-ended. For example, grand-tour questions were used to allow principals to identify factors influencing their decision to exclude (or not) a student from the LSA.

Each interview was audio-recorded and transcribed. Data were analyzed using a narrative analysis which wove together the structural, functional, thematic, and dialogic/performance components of narrative analysis (Parcell & Baker, 2017). The structural analysis considers the characteristics of the school and principal, whereas functional analysis examines the purpose or significance of what is being told. The thematic analysis identifies the common threads in each interview such as transparency, compassion, or inexperience and dialogic/performance focuses on the logistics of the interview, such as where the interview was conducted, who was involved in the interview, and any stakes involved in the sharing of information (Parcell, & Baker, 2017).

Findings

In this section, school profiles are provided and thematic findings are described. Findings are organized as (a) factors influencing principals' decision-making, (b) the team decision-making process, (c) exclusion practices, and (d) changing rates of exclusion.

School Profiles

Nine administrators from eight schools (School A, B, C, D, E, F, G, H) participated in this study. (The principal and vice-principal from School G participated in the interview together). Schools represented a range of grades and configurations: elementary (1), elementary and middle (3), middle (1), high (1), and K–12 (2) schools. All schools participated in provincial assessments, seven schools participated in international assessments (PISA), and six schools participated in national assessments (PCAP). Enrollments ranged from less than 200 to more than 500 students. Schools were located in all regions of the province, with more rural than urban schools represented. Four schools (D, E, F, G) were identified as serving communities with low socioeconomic status (SES), while School A was identified as having students with varied SES. Principals of the other three schools did not comment on the SES status of their communities. Four principals noted a range of community services, included snacks, meals, and clothing cupboards at schools. Table 1 presents a summary of school profiles.

Factors Influencing Principals' Decision Making

Principals considered several factors in making decisions to include or exclude students from LSAs, including exclusion criteria provided, students' education plans, educational supports and adaptations, students' progress on grade level outcomes, and other factors, such as community socioeconomics. All principals in this study referenced criteria provided by administering bodies and the province. Several principals in the study described in detail the differences between Individualized Educational Plans (IEPs) and Academic Learning Plans (ALPs). All principals referenced educational supports in the form of adaptations. Many noted the extent to which

Table 1

Demographic Characteristics of Schools

Principal	Grades	LSAs	Enrolled	Location
Principal A	K–12*	PISA, PCAP, Provincial	>350	urban
Principal B	10–12	PISA, Provincial	>400	rural
Principal C	7–9	PISA, PCAP, Provincial	>500	rural
Principal D	K–9	PISA, PCAP, Provincial	<200	rural
Vice Pr. D		as above		
Principal E	K–9	PISA, PCAP, Provincial	<250	rural
Principal F	K–6	Provincial	>200	urban
Principal G	K–12	PISA, PCAP, Provincial	>500	rural
Principal H	K–9**	PISA, PCAP, Provincial	>200	rural

*French School
**French Immersion Program

students on ALPS were working at grade levels, and almost half of the principals mentioned other factors, such as socioeconomics, in providing context for exclusion decisions.

Exclusion Criteria

Overall, principals were well aware of exclusion criteria on international and national assessments. However, there was some indication that exclusion criteria for provincial LSAs were in the process of being refined and poorly communicated and implemented. Recently, the province had published more detailed guidelines and criteria, including information about which adaptations would (and would not) be permitted during testing; however, these criteria were unevenly applied.

In particular, all principals in this study distinguished between students with IEPs and students with ALPs. Students with IEPs were characterized as having exceptionalities that dictated a modified curriculum but who nevertheless participated in daily school life to the extent possible. For example, one principal spoke of a non-verbal, autistic student whose IEP reflected a completely modified curriculum. Another principal commented, “Students on an IEP would be learning life skills; their entire day is different than a regular school day.” Overall, principals reported fewer students with IEPs and more students with ALPs, particularly in the elementary years. By contrast, students with ALPs were characterized in terms of their mastery (or not) of grade-level curricular outcomes. In general, principals perceived ALPs as historical documents, providing teachers with specific information about which foundational outcomes (primarily in math and literacy) that a child had not yet mastered. ALPs were described in terms of “reaching back” and helping students to achieve missing outcomes. Principals explained that students with ALPs varied widely and could range from a child missing several outcomes to a child who was two or more grade levels behind in foundational curricular outcomes. Depending upon the needs of each child, a range of supports in the form of adaptations might be offered to support the student. Due to wide variation in the needs and supports of children with ALPs, principals indicated the need for individualized decision-making regarding exclusion from LSAs.

Adaptations

In making decisions about whether to include or exclude students in an LSA, principals and their teams also considered the nature of adaptations used by students with ALPs and whether or not these adaptations were permitted to be used during LSAs. Guidelines existed regarding the use of adaptations with all three types of LSAs; however, these guidelines were not consistent among the LSAs. For example, participants noted that provincial guidelines had changed from year to year, and principals reported differing application of guidelines for adaptations from school to school. Two principals described adaptations as “legal” requirements on the one hand, but on the other hand, these same adaptations were then disallowed on an LSA. Other principals expressed frustration that adaptations, such as number lines or use of computers to type, were not permitted during the provincial assessments, particularly if the child had been accustomed to its use for some time. As one principal said, “So the entire year that child used a computer [to type] but in the provincial assessment she should have the computer but she wasn't allowed to use the computer.” Another principal noted, “A number of the adaptations are not permitted on the provincial assessment and that is where teachers get upset; we have been testing our students all semester and they have been allowed to have an adaptation.” Similarly, another principal said,

> The disconnect is sometimes not all of the adaptations are put into place in their provincial assessment and that's really frustrating for students and for teachers. Because we work so hard to have a kid learn to be successful with those adaptations. And sometimes, during the provincial assessments, they're taken away. And teachers feel horrible about that. And they don't like it at all, rightfully so.

Conversely, another principal noted that an adaptation, such as the use of a number line, was permitted. In the same vein, another principal reported that adaptations were permitted, especially if its use was "in their [ALP] plan" and had been documented previously, noting "if they were using voice to text and that was a documented adaptation that they have had for the entire year, then they did that." A third principal noted that the use of Chromebooks was a permitted adaptation. In sum, principals welcomed guidance from the province; however, criteria for the use of adaptations were unevenly applied.

Making Decisions: A Team Process

All principals in this study described themselves as the final decision maker in this process and held themselves accountable for decisions made; however, they did not make their decisions in isolation. All but one principal indicated that they would, at a minimum, make the decision in concert with the classroom teacher. As one principal put it, "Oh, the classroom teacher, definitely, yes. Their data is the strongest. They've got the best knowledge of the student." Three principals indicated that they would consult with a district administrator in the decision and a fourth principal described consulting with district curriculum specialists. Four principals detailed the members of a student services team who would weigh in on the decision, including such personnel as academic resource teachers, literacy coaches, and guidance counselors. Two of these four principals specified behaviour resource teachers as part of their teams. One principal described the province as the final decision maker, indicating that guidelines provided dictated the decisions to be made. Finally, almost all principals mentioned efforts to educate parents about the importance of LSAs and almost half of the administrators (four) indicated involving parents in the decision to include or exclude participation in LSAs. In sum, none of the principals saw themselves as making decisions by themselves. As one principal put it "But the teacher would never decide alone, I would never decide alone, our resource teacher would never decide alone. But as a team of three, we will decide and then often get permission as to why they should or shouldn't." Clearly, principals saw themselves as leading teams of decision-makers.

Grade Level Outcomes, Fairness, and Exclusion Practices

All administrators in this study confirmed that some students in their schools were working below current grade levels and principals across all grades discussed using grade-level curricular outcomes as a measure in making the decision to include or exclude a student from participating in LSAs. In general, principals considered whether or not students were operating two or three years below grade level or if students were operating somewhat below grade level. For example, if a student was missing selected foundational outcomes in math or literacy from the previous one or two years, they would be operating somewhat below grade level. Principals reported varying exclusion practices and often raised the issue of fairness in making exclusion decisions based on grade-level outcomes.

One principal noted that students who were two years or more behind grade level would have an IEP and would not participate in LSAs. The principal asked, "So why put a kid in that struggle when we know they are not going to achieve? They are not going to be able to write [the LSA]." However, this principal indicated that students with ALPs would write, noting "we try to get everybody to write the evaluation [LSA]." Another principal commented that students with IEPs would be excluded automatically from participation in LSAs due to the nature of their modified curriculum. This principal invoked the issue of fairness and asked,

> If they have not seen any of those outcomes, they should not be assessed on them. If they are in a completely modified program, they have never been exposed to those outcomes, so why would we assess them on the outcomes?

The principal added that students on ALPs would be included because "if they covered the material … at least the majority … if they have covered some of the grade 9 outcomes, they should be writing that assessment." Similarly, a third principal commented on students working "significantly below grade level" noting, "If you had a student in grade 6 who is working at grade 3 or 4 curricula in math, it would not be of benefit for that student to write that grade 6 math assessment." This principal went on to say that if the student were at grade level in reading and writing, they would participate in the writing assessment.

In addition to issues of fairness with grade-level outcomes, two principals alluded to the decision-making process as time and resource-intensive. An elementary school principal gave the example of grade-level outcomes in the early elementary years, noting students

> … who haven't done all of the grade three outcomes or been exposed to all of them, because we're still working on kindergarten and grade one to support them. So, it isn't fair for them to write an assessment that they haven't seen the work. So, then it became a case-by-case basis on looking at what is the specific programming; what reach-back have we done?

This principal went on to describe the school's decision-making process as "lengthy", especially in the case of students with ALPs. Likewise, another principal described beginning the decision-making process "early in the year" for students with ALPs. This K-12 principal reported,

> Usually, in January/February, we start to identify some kids that we would naturally determine would be excluded. So, students on an IEP, and then we start to kind of tease out those that are on an ALP. And we kind of determine how far back they're reaching. So, if they're reaching back three grades, yes, we would exclude them. If they're reaching back a year or two, that's when we have to start having conversations with our inclusive ed. consultant, with our student services team … It takes some time, sometimes to determine who you're going to be prepared to … [exclude].

Both of these principals referred to the time and care with which exclusion decisions were considered.

Last, three principals alluded to clearer-cut decision-making, even in the case of students with ALPs. The first principal estimated that 30% of students in the school were missing select outcomes. All students in this school wrote LSAs, even students with IEPs. Students who attended a transition program (which led to a certificate, not a high school diploma) would not participate in LSAs, nor would students with severe mental health issues (such as suicidal ideations). Similarly, another principal indicated that all students on ALPs would participate in LSAs, noting

"there are very few exceptions at our school." Although other principals indicated being knowledgeable about the impact of exclusion decisions, this principal explicitly made the connection to the quality of the data, stating: "because we want to make sure that we're going to be able to read the data accurately." Both of these principals indicated clear cut criteria and reported practices at their respective schools that suggested lower rates of exclusion.

Exclusion Rates

Overall, principals were knowledgeable about the reasons for holding exclusion rates to 5%; however, they held differing opinions about reasons for exclusion rates. In general, principals perceived their own schools' exclusion rates as low, although rates of exclusion throughout the province on various LSAs had exceeded recommended levels in recent years, which suggests evidence of what is known as the *Lake Wobegon effect* which is a tendency for most people, in this case, principals, to believe their school is above the provincial average (Lake Wobegon effect, n.d.).

A variety of factors were mentioned as influencing higher than recommended rates of exclusion. For example, one principal attributed higher rates to a national commitment to differentiation in education, noting that the practice "is creating high exclusion rates." The principal went on to comment, "By trying to educate every kid differently in classrooms, then we exclude many kids from writing common assessments." Two principals thought that high numbers of second language students or high numbers of transient students in urban schools might contribute to higher exclusion rates. In their own school, the latter principal noted that exclusion rates were highly cohort dependent and, therefore, would fluctuate from year to year.

Three of the principals reported low or very low rates of exclusion, depending upon the grade level. For example, one K-9 principal reported no exclusion in the lower grades and 9% exclusion in the upper grade. Another K-12 principal estimated that 1-2 students, at most, per grade level would be excluded and described exclusion rates as "very low." A third principal said their school kept to "lower" rates than the 5% maximum exclusion rate suggested. Each of the three high school principals and one middle school principal indicated that all students wrote the higher-stakes assessments in grade 9 (mathematics), grade 10 (language arts), and grade 11 (mathematics) because they "had to."

Several principals commented on their exclusion rates that had been higher but were now lower. For example, one principal reported exclusion rates that had dropped since the school had begun following revised regulations provided by the province, commenting, "When we know better, we do better." Another new principal commented that prior to their tenure, more students "were being excluded" and "exclusion rates were higher." However, this urban principal acknowledged that if the school were to follow the exclusion criteria to the letter, "almost 40% would be excluded." Another principal acknowledged that recent exclusion rates were higher and that current practice was to bring the rates down to below the recommended 5%. Last, another principal described their school's exclusion rate as adhering closely to the recommended rate of five to six percent.

Most principals referenced the support provided by newly revised provincial guidelines and described exclusion practices in their own schools. Most did not comment on a system-wide practice. However, one principal expressed frustration with an uneven application of exclusion criteria across the system. In acknowledging differing rates of exclusion from school to school, this principal commented, "I'm not a firm believer in LSAs. Don't get me wrong. But if we are

going to do them, let's be consistent [in how we exclude].” These and other quotes indicated both an appreciation for increased guidance from the province and a growing awareness of the need to apply consistently such guidance in order to garner an accurate perspective of how the system is working for students.

Discussion

This study confirms what other researchers reported related to the absence of clarity and governance when excluding students from LSAs (Brzyska, 2018; Eivers, 2010; Gamazo et al., 2019). In addition, we are able to extend the body of knowledge related to exclusion rates and identify several interconnecting factors influencing principals' decisions to include or exclude students from LSAs. These factors include awareness of student supports that are permitted on LSAs and excluding students who have IEPs or ALPs, which is interwoven with social promotion and compassion.

Principals were cognizant of guidelines for exclusion on the international and national LSAs but less so on provincial LSAs, which echoes previous research about loosely defined exclusion criteria (Braun et al., 2010; Brzyska, 2018; Gamazo et al., 2019; Prais, 2003; Wuttke, 2007). Specifically, there was uncertainty around what supports students could and could not use. For some students, their learning plan called for an adaptation in their learning, such as using a computer to type, voice-to-text technology, or a number line in mathematics. As we have found in our study, there was conflicting understanding of guidelines where one principal excluded a student from participating in a LSA because the student's education plan indicated the use of a computer to type but the principal believed this was not allowed on the provincial LSA and claimed a disconnect between practice and policy. This provides evidence that exclusion practices can vary from one principal to the next. There was some consensus as most principals permitted the use of computers for typing on the provincial LSA and believed all that was required was to note the adaptation received on the assessment paper. It is possible the absence of oversight in the administration of LSAs has allowed this difference in practice to occur. Further, it raises awareness of the need to align adaptations received during the school day and on LSAs and to communicate this alignment to principals to ensure that principals' decision-making is informed with current practice. It is important to note that as the number of students with special learning needs increases, it is likely that additional supports and adaptations to student learning will be required. These supports should be accommodated on LSAs, given the mandates for inclusive education (Joy & Murphy, 2012; Pollock & Winton, 2016).

Students with special learning needs typically have an IEP or, in the case of the province studied, an ALP. There was some consistency related to automatically excluding students with an IEP where five principals and one vice-principal indicated that students with IEPs were automatically excluded. However, three principals indicated they reviewed these students' files to determine whether students should or should not write the LSA. It is suspected that this understanding of IEPs would vary between provinces given that all Canadian jurisdictions have not introduced ALPs, thus in other provinces, students with IEPs would have a wider range of abilities, each needing individualized instruction. In the province we studied, IEPs were reserved for students with severe cognitive disabilities; although, two principals indicated that students who were socially promoted several grades also had an IEP. Thus, there does not appear to be a consistent practice of identifying who receives an IEP, who receives an ALP, and who participates in the LSAs.

In the province we studied, in most cases, students who fell short on meeting curriculum outcomes would be given an ALP. It is important to note that these students do not have intellectual disabilities but rather, for one reason or another, they did not achieve the curriculum outcomes in the prior grade(s) and were socially promoted to the next grade level. All principals confirmed that they had students operating below current grade level and one principal indicated that as many as 40% of her student population were operating below grade level. Subsequently, these students would be placed on an ALP. Deciding whether to include or exclude students with ALPs presented tension for principals, given students' wide range of abilities.

The threshold for excluding students appeared to be around two grade levels. For example, if a student was operating two or more grade levels behind their age-appropriate peers, it was likely that student would be excluded from writing the LSA. It would be amiss if we did not note that if students were operating two or more grade levels behind their peers, then despite promoting an age-grade correspondence, social promotion has reduced the number of achieved learning outcomes and the age-grade correspondence is likely to be lost when students enter secondary school where repetition of grades is the consequence for unsuccessful achievement (Leighton et al., 2019). In returning to our primary discussion, the threshold for excluding students who had been socially promoted was set at one or more grade levels behind for one principal. In summary, this study provides evidence of the entanglement between social promotion and excluding students from writing LSAs which can be connected to decision-making based on compassion.

Two principals described how they based decisions on compassion. For example, one principal described a student who had a large math gap because it called for abstract thinking and the student would shut down. Therefore, the principal excluded her from the LSA. Another principal described a similar situation and stated, "why put a kid in that struggle when we know they are not going to achieve. They are not going to be able to write [the LSA]." Both cases involved decisions based on compassion for students' wellbeing. Further, these decisions were more likely to occur in the primary and junior grades and exemplify Braun's et al. (2010) concerns about the creditability of LSAs in that low achieving students were more apt to be absent from the assessment. It is important to note that these students' abilities were not diagnosed as intellectual disabilities but rather as students who had fallen far behind their peers. The question as to why students had fallen so far behind their peers is another research study that warrants investigating.

Should all students write the assessment as called for by (Zumeta, 2015), or should we formally recognize the practice of social promotion and its implications on LSAs? In the case of the latter, it may be necessary to explicitly report the percent of students excluded from the LSA because they were operating below grade level. This delineation would eliminate the ambiguity in reporting the percent of students excluded from LSAs under the *non-participant* or *other* categories on the international and national LSAs and at the same time prevent exposing students to an assessment instrument they do not have the academic ability to complete (CMEC, 2018, 2019).

At this point, it is necessary to also raise implications of the shift towards inclusive education and the multiple mandates calling for inclusive practices and differentiated instruction on LSAs (Joy & Murphy, 2012; Pollock & Winton, 2016). As noted earlier, provincial and territorial law mandates that children with disabilities are entitled to receive accommodations to benefit from public education. The challenge here is that most children experiencing difficulties in schooling are not disabled but have learning difficulties which also requires some form of intervention. These interventions may be considered best practice rather than a legal obligation and the number of students requiring interventions appears to be growing (Learning Disabilities Association of

Ontario, 2018; National Center for Education Statistics, 2021). One principal believed that the practice of differentiating instruction through interventions was contributing to high exclusion rates, given that the multiple ways of differentiating instruction were not carried over when the student wrote LSAs. Thus, this principal was placed in a difficult position knowing that a student should otherwise write the LSA but at the same time, the principal believed it was not fair for the student to remove the supports they have received throughout the year.

When the interview focused on the school's exclusion rates, almost all principals were cognizant of the 5% benchmark set by provincial (as mentioned by principal C and E) and national regulatory bodies (OECD, 2017). Further, it appears that the gatekeeping purpose of high-stakes assessments results in lower exclusion rates and that principals' decision to exclude a student on such high-stakes assessments may be influenced by the purpose of the assessment. Based on the outcome of interviews with four principals, there appeared to be a provincial impetus to increase student participation on the lower-stakes LSAs, which likely influenced their decisions to reduce the number of students excluded from the LSAs. However, based on the findings from this study, there is still several misguided or misinformed principals along with the misalignment between inclusion practices and LSAs which aligns with the concerns expressed by Anders, et al. (2021) and Miller and Yan (2022).

Conclusion and Implications

In reflecting on this study, principals were faced with tension related to knowing whether to include or exclude a group of students who have fallen through the cracks in the education system, which is a decision that appears to be connected to the purpose or stake of the assessment. This group of students excluded from the LSAs have the necessary communication skills to write LSAs and they do not have severe cognitive disabilities. Instead, these students are, in most cases, products of social promotion that push students from one grade to the next without having achieved the learning outcomes. Principals were subsequently presented with two options. In the first option, all students should write LSAs to obtain an accurate picture of student learning and school accountability. In doing so, there would be many students writing an assessment who would have never been exposed to the types of items on the LSA. This practice would counter what principals know about good assessment practices, but stakeholders would have an accurate picture of student learning and an indicator of how well those responsible for educating children were doing their job. The second option would be to exclude this group of students who have fallen behind since it would not be fair and some may argue, inhumane to expose children to rigorous assessments they are not prepared to write. The outcome of this decision would result in an upward bias in achievement scores and the accountability framework of LSAs would be compromised thus raising the question, why bother with LSAs at all if the data is not accurate. We propose a third option which weaves together an understanding of the inappropriateness to include students in a LSA who are not in a position to write the assessment and, at the same time, accounts for their absence from the LSA. In this scenario, students who have fallen at least two grades behind their age-equivalent peers would be excluded from the LSA; however, the number of these non-writers would be documented and reported on LSA reports. Such a practice would provide an accurate account of student learning and not expose students to assessments they are not prepared to write. We recommend that this new approach to reporting student participation on LSAs should be piloted and validated.

References

Ahmed, A. Y., & Mihiretie, D. M. (2015). Primary school teachers and parents' views on automatic promotion practices and its implications for education quality. *International Journal of Educational Development, 43*, 90–99. https://doi.org/10.1016/j.ijedudev.2015.05.003

Albrecht, S. F. & Joles, C. (2003). Accountability and access to opportunity: Mutually exclusive tenets under a high-stakes testing mandate. *Preventing School Failure: Alternative Education for Children and Youth, 47*(2), 86–91. https://doi.org/10.1080/10459880309604435

Anastasiou, A. I., Papachristou, E. M., & Diakidoy, I. N. (2017). Parents' and teachers' views on the psychosocial adjustment of students with and without a history of early grade retention. *Educational Psychology in Practice, 33*(3), 294–307. https://doi.org/10.1080/02667363.2017.1306490

Anders, J., Has, S., Jerrim, J., Shure, N., & Zieger, L. (2021). Is Canada really an education superpower? The impact of non-participation on results from PISA 2015. *Educational Assessment, Evaluation and Accountability, 33*, 229–249. https://doi.org/10.1007/s11092-020-09329-5

Anderson, D., Lai, C. F., Alonzo, J., & Tindal, G. (2011). Examining a grade-level math CBM designed for persistently low-performing students. *Educational Assessment, 16*(1), 15–34. https://doi.org/10.1080/10627197.2011.551084

Anonymous (2020). *Student Enrollment K–12, 2019-2020*.

Ayers, L. P. (1909). *Laggards in our schools*. New York: Russell Sage Foundation.

Bédard, M. (2015, August 27). *Pedagogical autonomy and accountability: A recipe for improving academic results*. MEI. http://www.iedm.org/sites/default/files/pub_files/note0815_en.pdf

Braun, H., Zhang, J., & Vezzu, S. (2010). An investigation of bias in reports of the National Assessment of Educational Progress. *Educational Evaluation and Policy Analysis, 32*(1), 24–43. https://doi.org/10.3102/0162373709351137

Brzyska, B. (2018). Trends in exclusion rates for students with special educational needs within PISA. *Oxford Review of Education, 44*(5), 633–650. https://doi.org/10.1080/03054985.2018.1496907

Chataa, B. M., & Nkengbeza, D. (2019). Challenges faced by primary school teachers in implementing the automatic promotion policy at a school in the Zambezi region. *Creative Education, 10*(7), 1731–1744. https://doi.org/10.4236/ce.2019.107123

CMEC. (2011). *PCAP 2010: Report on the Pan-Canadian Assessment of Mathematics, Science, and Reading*. http://www.cmec.ca/publications/lists/publications/attachments/274/pcap2010.pdf

CMEC. (2014). *PCAP 2013: Report on the Pan-Canadian Assessment of Science, Reading, and Mathematics*. https://cmec.ca/Publications/Lists/Publications/Attachments/337/PCAP-2013-Public-Report-EN.pdf

CMEC. (2018). *PCAP 2016 Report on the Pan-Canadian Assessment of Reading, Mathematics, and Science*. https://www.cmec.ca/publications/lists/publications/attachments/381/pcap-2016-public-report-en.pdf

CMEC. (2019). *Measuring up: Canadian results of the OECD PISA 2018 study*. https://www.cmec.ca/Publications/Lists/Publications/Attachments/396/PISA2018_PublicReport_EN.pdf

Connor, D. M. (2018). *Teachers' perspectives of low student achievement within social promotion practices* [Unpublished doctoral dissertation]. Walden University. https://scholarworks.waldenu.edu/dissertations/5173/

Crepeau-Hobson, F. (2016). Grade retention or social promotion. *National Association of School Psychologists, 44*(7), 14–16. https://eric.ed.gov/?id=EJ1200566

D'Alessio, S., & Watkins, A. (2009). International comparisons of inclusive policy and practice: Are we talking about the same thing? *Research in Comparative and International Education, 4*(3), 233–249. https://doi.org/10.2304/rcie.2009.4.3.233

Doherty, K. (2004, September 21). Social promotion. *Education Week*.

http://www.edweek.org/ew/issues/social-promotion/
Eivers, E. (2010). PISA: Issues in implementation and interpretation. *The Irish Journal of Education, 38*, 94–118. https://www.jstor.org/stable/pdf/20789130.pdf
Elliott, J. L., & Thurlow, M. L. (1997). *Opening the door to educational reform: Understanding educational assessment and accountability*. https://files.eric.ed.gov/fulltext/ED412693.pdf
Emler, T. E., Zhao, Y, Deng, J., Yin, D., & Wang, Y. (2019). Side effects of large-scale assessments in education. *ECNU Review of Education, 2*(3), 279–296. https://doi.org/10.1177/2096531119878964
Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., & Rodríguez-Conde, M. J. (Eds.). (2019). *Assessment of students with special education needs in PISA: Participation guidelines and data treatment*. In Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'19). Association for Computing Machinery, 187–191. https://doi.org/10.1145/3362789.3362919
Graham, L. J., & Jahnukainen, M. T. (2011). Wherefore art you, inclusion? Analyzing the development of inclusive education in New South Wales, Alberta and Finland. *Journal of Education Policy, 26*(2), 263–288. https://doi.org/10.1080/02680939.2010.493230
Grek, S. (2009). Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy, 24*(1), 23–37. https://doi.org/10.1080/02680930802412669
Heydrich, J., Weinert, S., Nusser, L., Artelt, C., & Carstensen, C. H. (2013). Including students with special educational needs into large-scale assessments of competencies: Challenges and approaches within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online, 5*(2), 217–240. https://doi.org/10.25656/01:8431
Hinton, A. (2014). *Tracking student achievements in mathematics over time in English-language schools: Grade 3 (2006) to Grade 6 (2009) to Grade 9 (2012) cohort*. EQAO Research. Queen's Printer.
Hwang, S. H. J., & Cappella, E. (2018). Rethinking early elementary grade retention: Examining long-term academic and psychosocial outcomes. *Journal of Research on Educational Effectiveness, 11*(4), 559–587. https://doi.org/10.1080/19345747.2018.1496500
Joy, R., & Murphy, E. (2012). The inclusion of children with special educational needs in an intensive French as a second-language program: From theory to practice. *Canadian Journal of Education / Revue canadienne de l'éducation, 35*(1), 102–119. https://journals.sfu.ca/cje/index.php/cje-rce/article/view/712
Klieger, A. (2016). Principals and teachers: Different perceptions of large-scale assessment. *International Journal of Educational Research, 75*, 134–145. https://doi.org/10.1016/j.ijer.2015.11.006
Klinger, D., Deluca, C., & Miller, T. (2008). The evolving culture of large-scale assessment in Canada. *Canadian Journal of Educational Administration and Policy, 76*, 1–34. https://eric.ed.gov/?id=EJ807064
Lake Wobegon effect. (n.d.). *Oxford Reference*. https://www.oxfordreference.com/view/10.1093/oi/authority.20110810105237549
Learning Disabilities Association of Ontario. (2018, February). *Learning Disabilities Statistics*. http://www.ldao.ca/introduction-to-ldsadhd/articles/about-lds/learning-disabilities-statistics/
Leighton, M., Souza, P., & Straub, S. (2019). Social promotion in primary school: Effects on grade progression. *Brazilian Review of Econometrics, 39*(1), 1–33. https://ideas.repec.org/a/sbe/breart/v39y2019i1a78513.html
Leishman, A. M. (2015). *Accommodating students on classroom and large-scale assessments: Teachers' perspectives and validity issues* [Unpublished master's thesis]. Lakehead University. http://knowledgecommons.lakeheadu.ca/handle/2453/728
LeRoy, B. W., Samuel, P., DeLuca, M., & Evans, P. (2019). Students with special educational needs within PISA. *Assessment in Education: Principles, Policy & Practice, 26*(4), 386–396. https://doi.org/10.1080/0969594X.2017.1421523

McGrew, K. S., Thurlow, M. L., Shriner, J. G., & Spiegel, A. N. (1992). *Inclusion of students with disabilities in national and state data collection programs*. http://files.eric.ed.gov/fulltext/ED347769.pdf

McMahon, T. (2018). Despite our best intention: Students relate how social promotion hurt them and what changes they believe will help them. *Interchange, 49*, 499–519. https://doi.org/10.1007/s10780-018-9344-3

Miller, T. (2013). Using large-scale assessment results to determine student grades. Canadian, *Journal of Education / Revue canadienne de l'éducation, 36*(3), 317–353. http://journals.sfu.ca/cje/index.php/cje-rce/article/view/939/1609

Miller, T., & Yan, Y. (2022). Factors influencing the accuracy of Canada's large-scale assessment data: Policies and practices of exclusion, absenteeism, and social promotion. *Educational Assessment, Evaluation and Accountability*. https://doi.org/10.1007/s11092-022-09381-3

Murphy, D. (2014). Issues with PISA's use of its data in the context of international education policy convergence. *Policy Futures in Education, 12*(7), 893–916. https://doi.org/10.2304/pfie.2014.12.7.893

National Center for Education Statistics. (2021 May). *Students with disabilities*. https://nces.ed.gov/programs/coe/indicator_cgg.asp

Newton, P., Tunison, S., & Viczko, M. (2010). The school principal's role in large-scale assessment. *Canadian Journal of Educational Administration and Policy, 105*. https://journalhosting.ucalgary.ca/index.php/cjeap/index

Nichols, S., L., Glass, G., V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives, 14*(1). https://doi.org/10.14507/epaa.v14n1.2006

OECD. (2011). *School autonomy and accountability: Are they related to student performance? PISA in Focus, 9*. https://www.oecd.org/pisa/pisaproducts/pisainfocus/48910490.pdf

OECD. (2017). *PISA 2015 technical report*. https://www.oecd.org/pisa/data/2015-technical-report/PISA2015_TechRep_Final.pdf

Parcell, E. S., & Baker, B. M. A. (2017). Narrative analysis. In M. Allen (Ed.), *The SAGE encyclopedia of communication research methods* (pp. 1069–1072). SAGE Publications. https://doi.org/10.4135/9781483381411.n368

Pollock, K., Wang, F., & Hauseman, C. (2017). Complexity and volume: An inquiry into factors that drive principals' work. In K. Leithwood, S. Jingping, & K. Pollock (Eds.), *How school leaders contribute to student success: The four paths framework*. (pp. 209-238). Springer International Publishing.

Pollock, K., & Winton, S. (2016). Juggling multiple accountability systems: How three principals manage these tensions in Ontario, Canada. *Educational Assessment, Evaluation and Accountability, 28*, 323–345. https://doi.org/10.1007/s11092-015-9224-7

Prais, S. J. (2003). Cautions on OECD's recent educational survey (PISA). *Oxford Review of Education, 29*(2), 139–263. https://doi.org/10.1080/0305498032000080657

Prytula, M., Noonan, B., & Hellsten, L. (2013). Toward instructional leadership: Principals' perceptions of large-scale assessment in schools. *Canadian Journal of Educational Administration and Policy, 140*, 1–30. https://cjc-rcc.ucalgary.ca/index.php/cjeap/article/view/42840

Renihan, P., & Noonan, B. (2012). Principals as assessment leaders in rural schools. *Rural Education, 33*(3), 1–8. https://doi.org/10.35608/ruraled.v33i3.407

Rogers, W. T. (1993). Principles for fair student assessment practices for education in Canada. *Canadian Journal of School Psychology, 9*(1), 110–127. https://doi.org/10.1177/082957358500900111

Schleicher, A. (2009). Securing quality and equity in education: Lessons from PISA. *Prospects, 39*, 251–263. https://doi.org/10.1007/s11125-009-9126-x

Schuelka, M. J. (2013). Excluding students with disabilities from the culture of achievement: the case of the TIMSS, PIRLS, and PISA. *Journal of Education Policy, 28*(2), 216–230.

https://doi.org/10.1080/02680939.2012.708789
Schulze, R. (2014). *School principal leadership and special education knowledge* [Unpublished doctoral dissertation]. University of Massachusetts. https://doi.org/10.7275/hr4c-be46
Sealy, K. M., Perry, S. M., & DeNicola, T. C. (2016). *Relationships and predictors of principal job satisfaction across multiple countries: A study using TALIS 2013 and PISA 2012*. https://doi.org/10.2139/ssrn.2717056
Smith, E., & Douglas, G. (2014). Special educational needs, disability and school accountability: An international perspective. *International Journal of Inclusive Education, 18*(5), 443–458. https://doi.org/10.1080/13603116.2013.788222
Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large-scale assessments*. (Synthesis Report 44). University of Minnesota, National Center on Educational Outcomes. https://nceo.umn.edu/docs/OnlinePubs/Synth44.pdf
Thurlow, M. L. (1995). *Inclusion of transition-age students with disabilities in large-scale assessments*. University of Minnesota, National Transition Network. https://files.eric.ed.gov/fulltext/ED385989.pdf
Timmons, V. (2008). Towards inclusive education in Canada. In: Barton L., Armstrong F. (Eds.). *Policy, experience and change: Cross-cultural reflections on inclusive education* (pp. 133–146). Springer.
Volante, L. & Ben Jaafar, S. (2008). Educational assessment in Canada. *Assessment in Education: Principles, Policy, and Practice, 15*(2), 201–210. https://doi.org/10.1080/09695940802164226
von der Embse, N. P., Pendergast, L. L., Segool, N., Saeki, E., & Ryan, S. (2016). The influence of test-based accountability policies on school climate and teacher stress across four states. *Teaching and Teacher Education, 59*, 492–502. http://dx.doi.org/10.1016/j.tate.2016.07.013
Wanzek, J., Otaiba, S. A., Schatschneider, C., Donegan, R, E., Rivas, E., Jones, F., & Petscher, Y. (2020). Intensive intervention for upper elementary students with severe reading comprehension difficulties. *Journal of Research on Educational Effectiveness, 13*(3). 408-429. http://doi.org/10.1080/19345747.2019.1710886
Wuttke, J. (2007). Uncertainties and bias in PISA. In Hopmann, S. T., Brinek, G., & Retzl, M. (Eds.), *Does PISA keep what it promises?* (pp. 265–294). LIT.
Yan, Y. (2016). Public knowledge and perceptions of large-scale assessments [Unpublished master's thesis]. University of Prince Edward Island.
Ysseldyke, J., & Thurlow, M. (1994). *Guidelines for inclusion of students with disabilities in large-scale assessments*. NCEO Policy Directions, Number 1. https://files.eric.ed.gov/fulltext/ED372560.pdf
Zumeta, R. (2015). Implementing intensive intervention: How do we get there from here? *Remedial and Special Education, 36*(2), 83–88. https://doi.org/10.1177/0741932514558935

Dr. Tess Miller is a Professor at the University of Prince Edward Island and a Credentialed Evaluator with the Canadian Evaluation Society, PEI Chapter. She teaches graduate level courses in program evaluation, quantitative statistics, and mathematics and assessment courses at the bachelor level. Her research focuses on mathematics and assessment as well as in developing quantitative instruments for gathering data in a variety of fields.

Dr. Anne Marie FitzGerald is an Assistant Professor in the Faculty of Education at the University of Prince Edward Island. She is a former second language teacher, curriculum developer, and elementary school principal and has worked in diverse educational settings in Canada, Colombia, Kuwait, and the US. She has presented papers at conferences such as UCEA and AERA on Latino family engagement; community, school, and university partnerships; and managed school choice. Her research interests

include teacher preparation, culturally relevant orientation for international students, and principals' decision-making processes.

Han Xu is a PhD student at Queen's University. Her research interests include internationalization of higher education, recruitment and admission, international students, and language programs. Han is a firm proponent of sustainable internationalization based on equity, diversity, and academic excellence.