



OECD Education Working Papers No. 68

A Literature Review of School Practices to Overcome School Failure

Brenton Faubert

https://dx.doi.org/10.1787/5k9flcwwv9tk-en



Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

08-Feb-2012

English - Or. English

DIRECTORATE FOR EDUCATION

EDU/WKP(2012)5
Unclassified

A LITERATURE REVIEW OF SCHOOL PRACTICES TO OVERCOME SCHOOL FAILURE

Directorate for Education Working Paper N°68

By Brenton Faubert

January 2012

This working paper was prepared as part of the OECD thematic review Overcoming School Failure: Policies that Work, www.oecd.org/edu/equity. The project provides evidence on the policies that are effective to reduce school failure by improving low attainment and reducing dropout, and proactively supports countries in promoting reform. The project builds on the conceptual framework developed in the OECD's No More Failures: Ten Steps to Equity in Education (2007). Austria, Canada (Manitoba, Ontario, Québec and Yukon), Czech Republic, France, Greece, Ireland, Netherlands, Spain and Sweden took part in this project.
This working paper is part of a series of papers prepared for the thematic review Overcoming School Failure: Policies that Work covering the topics of policies to reduce dropout and in-school practices to reduce school failure. These report have been used as background material for the final comparative report Equity and Quality in Education: Supporting Disadvantaged Schools and Students (OECD, 2012), which gives evidence on the policy levers that can help overcome school failure and reduce inequities in OECD education systems. It focuses on the reasons why investing in overcoming school failure -early and up to upper secondary- pays off, on alternatives to specific system level policies that are currently hindering equity, and on the actions to be taken at school level, in particular in low performing disadvantaged schools.

Beatriz Pont, + 33 (0) 145 24 18 24, beatriz.pont@oecd.org

JT03315648

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

TABLE OF CONTENTS

Introduction....3
1. Overview of the literature....4
2. Areas of policy and practice to reduce school failure by improving equity in schools and classrooms....5
3. Challenges and considerations in the design and implementation....21
4. Connections across policies and practices....24
5. Conclusion....25
REFERENCES....27
ANNEX A: RESEARCH QUESTIONS AND TERMS OF REFERENCE....31
ANNEX B: SEARCH METHDOLOGY....33
ANNEX C: HATTIE'S EFFECT SIZE BAROMETER....34

Introduction

The age/grade model of classroom and school organization emerged in the mid-19th century and has since become the standard approach to schooling across Organization for Economic Cooperation and Development (OECD) countries. Schools are composed of classes (and classrooms) of age/grade students. Students advance through grades, generally associated with an age, and classes are organized to deliver a grade of instruction. Promotion in the age/grade model is not guaranteed, meaning students can also fail to advance. One indication that failure is an outcome can be found in the OECD's 2010 *Education at a Glance,* which reports that 22 of 26 OECD member countries had first time upper secondary graduation rates above 70% and only a small number had rates of 90% or greater (OECD, 2010b). According to OECD statistics, this means that (on average) 20% of students at the end of four years of secondary school in nearly all OECD countries fail or opt to leave. In other words, 20% of students have not acquired the skills, knowledge or credits necessary to graduate from secondary education.

The outcome of *failure* in the age/grade model has served as a method of sorting students in educational systems (OECD, 2009). Further, the sorting of students through failing or advancing in educational systems has long been accepted as a satisfactory educational model. Since the 1960s, however, the view that student failure is necessary or inevitable has come under increasing scrutiny. An emerging viewpoint across OECD countries is that education systems must provide a successful educational outcome for all students. Increasingly, it is no longer seen as adequate to provide equal access to the same "one size fits all" educational opportunity. More and more, the onus is shifting to providing education that promotes equity by recognizing and meeting different educational needs. Importantly, this growing view also reflects a different understanding of the causes of student failure. **The idea that students fail because of their own personal shortcomings (academic or otherwise) is being superseded by the idea of school failure. The cause of – and the responsibility for – students failure is now seen as deficient or inadequate provision of education by schools, and by extension, school systems. More specifically, it is the failure of schools to provide education appropriate to different needs that leads students to fail. In this way school failure is, therefore, also an issue of equity.**

This shift in thinking about failure can be seen in both academic and policy circles. The emergent view holds that it is no longer acceptable that millions of students fail in school, especially given how failure may negatively affect their capacity to succeed in their personal lives, the working world, and, more broadly, society. Reorienting educational systems towards the goal of promoting equity has been advanced as the necessary solution to, and redress of, the student failure that results from an age/grade – class/school model. The focus on ensuring equality of opportunity to education, which underpins the age/grade model, has come at the expense of not attending to difference, whether systemic in nature, as with socio-economic status, or specific to the individual, where many variables can affect ability and need.

The OECD takes the position that new policies and practices can revamp educational systems and the classes and schools which the systems are built upon so that the effects of difference on student achievement are counter-balanced and, ideally effectively neutralized. In 1998, the OECD released a report titled *Overcoming Failure at School* which calls for shifts in strategic educational policies to address the needs of disadvantaged groups aimed at "preventing and overcoming the learning difficulties of all young people, as part of a larger aim to attain higher, required levels of knowledge, skills and competencies" (OECD, 1998). This work led to the OECD 2007 report *No More Failures: Ten steps to equity in education*, with policy suggestions aimed primarily at the state and district levels to improve equity, and thereby reduce school failure (OECD, 2007).

To provide further assistance for member countries to achieve real improvement in reducing failure, the project *Overcoming School Failure: Policies that Work* (hereafter referred to as *Project*) was

established in 2009. Through this Project, the OECD plans to) identify and analyze initiatives that successfully tackle school failure, and support countries in implementing policies and practices in overcoming school failure (OECD, 2010) at the state, district, school and classroom levels.

This report is one of a series of literature reviews commissioned by the OECD in support of the third goal of the Project: supporting countries with the implementation of policies and practices in overcoming school failure. To this end, this present review examines English literature on in-school practices for overcoming school failure. The purpose is to answer two general questions: 1) what are the policies and practices that help reduce school failure and improve equity? And 2) what is the empirical evidence of their impact on reducing school failure and improving equity?

1. Overview of the literature

Key terms and definitions

Failure

This paper is concerned with *school failure*, understood as the failure of schools and the school system to provide the appropriate level of, and adequately defined services for, all students to be successful. In this definition, school failure includes all policy and practice within schools and classrooms (i.e., pedagogy, school leadership, professional learning, etc.). This definition stands in contrast to the more conventional view of failure as an outcome for students when they do not acquire the skills, knowledge or credits expected or required to advance in the age/grade model. The definition used here embeds a different causal relationship in that it is the failure of schools which results in the failure of students. Following from this, the focus for redress and remedy is the school, rather than the student. The policies and practices discussed in this paper are ones shown by empirical research to address school failure. By improving the equitable provision of services at the school and classroom levels, they aim is to reduce school failure, and thereby, student failure.

Equity

As has been suggested, there is a strong relationship between school failure and the issue of equity in education. The provision of needs-sensitive and differentiated services is intended to address inequity so that students outcomes are attributable only to their effort and cannot be explained by factors known to be associated with inequity such as socio-economic status, gender, race, ethnicity or disability (Levin, 2003; Wobmann, 2008). This approach to improving equity has two elements. The first is *inclusion* which means ensuring a basic standard of education for all. The second is *fairness*, which means ensuring that neither systemic inequity nor disability are obstacles to achieving educational potential (OECD, 2007). This view of equity is consistent with the shift in education politics in OECD countries over the past half-century (OECD 1998; 2007). Increasingly the concern is to minimize the effects of inequity through reforms in policy and practice.

Policy and practice

Policy can be thought of as statements of vision and mandate which provide guidance for process and practice. In examining the impact and effectiveness of policy reform, key questions to be asked include: *who* is the policy intended to effect? *Why* is the policy in place? *How* will it be implemented? In *what* ways? *When* and *where* the policy will have impact?

A practice or set of practices can be thought of as an intervention, guided by policy, to address specific issues. Practices vary significantly in their intentionality, organization and coherence. At one end of the spectrum are well-defined and implemented programs. At the other are changes in how things are

done that can be interpreted to have intent. In this literature review, the policies and practices considered have two common features: 1) they are intended to reduce school failure, directly or indirectly, by improving equity, and 2) they have been subject to empirical evaluation.

Methodology

Since the literature that addresses some aspect of school failure is very large, parameters had to be set. Included was empirical research on policies and practices that reduce school failure by improving equity at the school and classroom levels. By the terms of reference of this assignment, topics excluded from this review include system-level policies and specific policies and practices on: special education learning needs, English language learners or other specific student groups. In addition, empirical studies on policies and practices concerning chartered schools, religious schools, private schools, bilingual schools were omitted. Finally, specific empirical studies recommending policies and practices related to literacy, numeracy, and technology were excluded because each of these topics warrants its own comprehensive literature review. A complete description of the methodology is provided in Appendix B.

The remit of this review was to assess the literature on empirically supported policies and practices for effectiveness in reducing school failure. This was done by assessing the design of the study and its methodology for the degree of systemacy in the research process and rigour. In some cases, studies cited effect size, a measure of the strength between two variables. For the purposes of this review, a policy or practice initiative was considered effective if the effect size was $d = 0.40$ or greater (Hattie, 2009). See Appendix C for the supporting rational behind Hattie's $d = 0.40$ hinge point. Though not excluded from the study, initiatives with effect sizes below $d = 0.40$ are considered to be less effective.

2. Areas of policy and practice to reduce school failure by improving equity in schools and classrooms

The policy and practices discussed in this part of the review are organized under thematic headings. Each thematic section addresses a) what policy or practice is under review, why it is important in reducing school failure, and how it reduces school failure. This is followed by b) a critical commentary of the literature reviewed, the strength of the supporting evidence, and the key points to be taken. Where possible educational actors at the classroom (teacher, student), school (teacher, principal, parent) or district level (principal, district administrator) having influence in the realization of the policy or practice are also identified. Section 2 finishes with a summary overview of policies and practices aimed at addressing school failure and improving equity.

2.1 Pedagogy

In this review *pedagogy* is understood to refer to teacher approaches in instruction, assessment and curriculum. Specifically, this section is concerned with what content gets presented in the classroom, what methods are used for content delivery, and what methods for assessment are developed. We are concerned with pedagogy because the research indicates that if curriculum is made relevant to the day-to-day lived experiences of students, then students are more engaged in their learning. If teachers are equipped with direct and student-oriented instructional approaches, then teachers are better able to provide instructional support. If the assessment tool evaluates both factual and deeper conceptual understanding, includes regular feedback that is clear and task specific, and valid conclusions can be drawn, then any student is able to advance in their learning.

How does pedagogy work to reduce school failure? By establishing a policy and practice context that supports the effective use of curriculum, instruction and assessment, teacher pedagogy will improve. This will result in the more equitable provision of educational services. A number of empirically assessed

policies and practices were identified in the review related to instruction, assessment and curriculum. The most promising of these are discussed here, beginning with instruction.

2.1.1 Instruction

Schools should set guidelines to promote the use of direct and student-oriented instruction. Direct instruction is instruction built around problems with clear, correct answers that can be learned quickly. Student-centered instruction is associated with the teacher facilitating students' own inquiry by allowing students time to find solutions to problems on their own before the teacher demonstrates how a problem is solved (OECD, 2009; Rowe, 2007; Hoad et al., 2007). Some Western countries, including Australia, Canada, the United States, Finland, Austria, Iceland, and Norway have a stronger preference for student-oriented approaches. Eastern and southern European countries (Bulgaria, Poland, the Slovak Republic, Italy, Portugal, and Spain) broadly endorse the two, with a moderate preference for student-oriented approaches (OECD, 2009; Rowe, 2007). There is no consensus in the literature on which approach is superior; however an over-reliance on either approach is not recommended (OECD, 2008b; Rowe, 2007; Waxman et al., 2001).

Several direct and student-oriented instructional practices were identified and supported by empirical evidence to improve teacher pedagogy. These instructional approaches are discussed below:

- Space learning over time (Direct instruction):

This practice is founded on the idea that students will remember content better when they are exposed to it on more than one occasion. Five experimental classroom studies and hundreds of laboratory experiments have supported this hypothesis (Pashler et al., 2007). Implementing this practice is straightforward. By way of example, the classroom teacher can organize course content in such a way to review important curriculum materials, weeks (and even months) later. The format of the review can take many forms, including an assignment or short-answer questions covering material already presented.

- Offer a worked-out problem followed by a related unsolved problem (Direct instruction):

Offering a worked-out problem followed by a related unsolved problem models for the student how-to successfully complete the exercise (Hattie, 2009; Pashler et al., 2007). Worked-out examples should include a problem statement and the appropriate steps to solve the problem. This instructional approach is criticized in the literature for being lenient on students but its supporters argue that it is better to focus students' attention on the process of answering the question (building skill and confidence) than frustrating students who may not be ready to solve the problem on their own (Hattie, 2009).

- Culturally responsive instruction (Student-oriented instruction):

Culturally responsive instruction requires that the teacher value student concerns, needs and realities (i.e., family, community). It is hypothesized that if the teacher works from the existing knowledge of the student, a shared cultural understanding will develop that can serve as a bridge (or scaffold) to guide instruction. For example, teachers attuned to cultural differences in how a student communicates could more effectively draw out what the student knows and better communicate the purpose and goals of the lesson (OECD, 2008c; 2010c; Klinger & Edwards, 2006; Waxman et al., 2001).

- Concrete and abstract representations of concepts (Direct instruction)

Providing both concrete and abstract representations of concepts can lead to more flexible learning on the part of the student. One possible approach is called *concreteness fading*, or using a concrete

representation to introduce a concept or principles, and then systematically replace relevant components of the concrete representation with abstract representations (Pashler et al., 2007). A practical example could include introducing the concept of fractions by cutting a pie into six slices and demonstrating how much of the pie would be gone if 1/6 of a slice was given to a friend. The lesson could then be taught again in a more abstract form, replacing the pie with the cut out of a circle cut into six equal triangles (like a pie). The teacher can then demonstrate to students that even though the object may change, the concept of fractions remains the same.

Classroom experiments and quasi-experimental evidence also supports the use of *concept mapping* – a process by which the teacher introduces a new and abstract concept first by summarizing the main points of the lesson, and then identifying and synthesizing the major ideas, themes and interrelationships between the concrete and abstract parts (*d = 0.57*) (Hattie, 2009; Pashler et al., 2007). Concept mapping could be used for instance when teaching a lesson on the concept of time. The teacher could list all the main points of the lesson – what is time? Why do we have time? What is the purpose of the clock? The minute hand? The hour hand? And the difference between the 12 hour and 24 hour clock? The teacher could then proceed to draw relationships between the abstract and concrete lesson points with learning tools (manipulatives), including a clock, a minute/second timer, a linear graph of the passage of time on the board, or the use of a sundial.

An important reminder for educators when using any concrete/abstract approach is that teachers must play a direct role in drawing the connections for students between the concrete and abstract representations of the lesson. The expected outcome for students is better comprehension of the main points of the lesson (particularly for struggling students) and superior application of the principles of the lesson in cross-over learning situations (Pashler et al., 2007).

- Deep-questions (Direct instruction)

After students have mastered the factual content of a particular topic, teachers can ask deep or higher order questions that challenge students' understanding of the concept, including "what caused X, how did X occur, what if, what-if-not, how does X compare to Y, what is the evidence for X, and why is X important" (Pashler et al., 2007, p. 30). The practice of asking higher-order questions is supported by strong evidence, including dozens of experimental studies and laboratory experiments. Asking low and higher order questions can be done in any lesson. For example, a teacher may introduce the concept of natural disasters, say earthquakes. The teacher could begin the lesson by asking simpler questions (i.e., "Can anyone recall a recent example of an earthquake?") and then proceed to higher order questions by asking to theorize on what might cause an earthquake? How an earth quake compares to an aftershock? The teacher could then probe deeper asking students to connect earthquakes to the creation of other natural hazards, i.e., tsunamis, or connect earthquakes to the concept of tectonic plates. Teachers and students may need to have modelled how-to ask and answer deep questions to achieve full-effect. Hattie's meta-analyses found that teachers asked 300-400 questions per day but the majority of these questions were lower-order questions which typically refer to factual or short answer questions such as, what is the date today?

- Students working in groups (Student-oriented instruction)

Arranging students in particular group structures can lead to both improved student engagement in the lesson and improved student achievement. Educators should pay attention to the group structure because not all arrangements are equally effective. For example, *ability grouping* refers to teachers placing students within the same class into groups of similar ability. In the case of ability grouping, the evidence suggest this structure does **not** lead to measurable positive effects with low (d = 0.37), medium, (d = 0.19) or high ability students (d = 0.28) (Hattie, 2009).

A second grouping structure, *cooperative grouping*, involves the teacher placing students within the same class in a small, temporary grouping of students with mixed abilities and students are given tasks that require them to work cooperatively and rely on each other's skills in order to complete the task successfully. Cooperative grouping is **hypothesized** in the literature to improve cognitive outcomes (Waxman et al., 2001) but the research does not provide empirical evidence explaining *how* or under *what* conditions this happens. For example, how do teachers generate high expectations for the entire group? Adopt curriculum challenging for all students? Encourage effective sharing of learning among peers (Leithwood et al., 2004)? This absence in empirical evidence is a serious limitation to student grouping strategies because we know that students need to be taught how to work effectively in groups of any type or size (Hattie, 2009; Sebba, 2006). In other words, placing students in cooperative groups will not necessarily lead to improved student achievement and, possibly, could lead to negative outcomes for some if one or a small number of individuals complete all the work while the remaining group members contribute minimally or not at all.

2.1.2 Assessment

Assessment practices were also identified as key to reducing school failure. For students to benefit, it is recommended that schools set guidelines that promote the combined use of summative and formative assessment. Summative assessment summarizes a student's achievement at a particular moment in time. Formative assessment evaluates student progress in their learning relative to their learning goals and the information is used to set actionable next steps (Harlen, 2006).[1] It is noteworthy that there are not different kinds of assessment tools linked to formative and summative assessment; rather what matters is how the assessment is used.

The use of summative and formative assessment is different across OECD countries. For example, in England, Germany, Mexico and the United States, policy and practice tradition has led to a mostly summative approach in assessing students, using varied combinations of teacher and externally set tests. In France, teachers have no role in assessing the student summatively and formative assessment is largely integrated into pedagogy (Black & Wiliam, 2005). In all countries, assessment practices are influenced as much by the beliefs of educators as what the evidence has to say (Black & Wiliam, 2005; 2006). The importance of considering the culture of a school, country or jurisdiction in the take-up and of policies and practices is discussed in detail in Section 3.

The review of the literature identified a number of assessment practices that have been empirically assessed to be effective. The most promising of these practices are discussed below:

- Assessment followed by feedback

Assessment (summative or formative) followed by feedback stood out in the review as a practice with great potential to impact student achievement (d = 0.73) (Hattie, 2009). Feedback can be thought of as an exchange of information between the teacher and student concerning the student's performance or understanding. The review found the following to be characteristic of useful feedback: specific to the task or the knowledge gap, timely, regular, well-formatted, clear, constructive in tone, makes transparent the learning process, avoids comments on behaviour and provides actionable steps for students to move-forward with on their own (Black & William, 2006; Hamilton et al., 2009; Hattie, 2009; OECD, 2008c; Sebba 2006).

- Formative use of summative assessments

[1] The terms summative and formative assessments are used interchangeably in the literature with the terms *assessment of learning* and *assessment for learning* (Black & Wiliam, 2006; Earl, 2003; Harlen, 2006).

Teachers and students are encouraged to make formative use of summative assessments. A practical example is provided in Pashler et al. (2007). In preparation for a summative assessment, students should review materials then take a break (i.e., hour, day or week). After the break, students should then attempt to generate keywords and definitions and summarize the main points of what they reviewed (as they would for the summative test). When students check their responses against the answers, they become aware of what they actually know and can focus their study time accordingly (Black & Wiliam, 2006; OECD, 2005; Pashler et al., 2007).

- Formative evaluation of teacher learning programs

Teachers should receive training on how-to-do formative evaluation of their own learning programs. Hattie (2009) found that supporting teachers in effectively assessing how well they are moving forward with their learning programs led to improved student achievement across age, duration of treatment, and frequency of measurement $d = 0.90$ (Hattie, 2009). For optimal effect, teachers are encouraged to use of multiple sources of data (i.e., data collected from classroom assessment, school level data) when assessing their programs because teacher self-assessment alone will not produce the desired effect. If implemented appropriately, regular and formative evaluation of teacher programs should result in better informed decision making when teachers adjust their pedagogical approach and, ultimately, improve student achievement.

- Construct definition & interpretation

Educators must give more attention to the *construct* of an assessment as a means of advancing equity in schools and classrooms (Pullin, 2008; Wiliam, 2010). In Wiliam (2010), the point is made that drawing valid inferences from assessment is both a technical matter (design and implementation of procedures) and a matter of construct definition and interpretation.

In simple terms, the author argues that the assessment of students begins with consideration of the *construct* ("the ability or human characteristic required to perform a task successfully") followed by *construct definition* ("defining the ability or characteristic to be assessed") then the technical step of designing the assessment in such a way that will yield evidence that support inferences about the construct and finally, *construct interpretation* ("the different views about the kinds of interpretations that can be made from assessment outcomes") (Wiliam, 2010, p. 260, 278). This last point regarding the "right" inferences to draw is not a technical one but one of values, and so the most likely to lead to tension (Wiliam, 2010, p. 278). These may sound like academic concerns distant from everyday practice in classrooms but this is not the case. When teachers assess students, they are essentially making claims about student performance. Through assessment, we "know" if a student has or has not mastered ability.

In his work, Wiliam invites the reader to consider the issue of assessment in history. One teacher may believe testing factual knowledge is important while another may believe that assessing a student's ability to weigh and interpret historical evidence and arguments is more important. Typically this difference in opinion would impact how each teacher develops their assessment. Multiple choice is reasonable to draw valid inferences about knowledge of important historical dates and facts. Open ended question would better suit assessment of student ability to examine and interpret historical evidence (Wiliam, 2010). Which approach is better? As Wiliam argues, the first step is not to begin with the beliefs of the teacher but instead begin by defining the construct the teacher is looking to assess. Once the construct has been considered and then defined, the technical step requires the instructor choose the most appropriate assessment procedures, that is a method (or combination of methods) that will minimize the potential of any assessment bias and also yield evidence to support inferences about the construct according to how it was defined. Debate over whether or not the student "knows" the construct will be better off having reached agreement on what is the construct, its definition and means of assessment early on.

2.1.3 Curriculum

Attention was also given to school policies and practices dealing with curriculum. Like instruction and assessment, how curriculum is organized varies across OECD countries. Some countries allow for greater regional/local responsibility in determining curriculum (Australia, Germany), while others opt for a more centralized curriculum (England, France, Japan, Mexico, New Zealand, Sweden). Some jurisdictions set a required number of requisite courses or clear pathways (France, Sweden), while others allow for more student choice (England, Germany, the United States) (Black & Wiliam, 2005). The variability in curriculum authority and choice across countries makes it difficult to provide clear policy and practice direction on curriculum at the school level but there are noteworthy trends.

Curriculum should be common and set high expectations for all, be linked to clear learning goals, and be connected to pathways that lead into tertiary education or employment. As much as possible, classroom activities should develop student knowledge of real world (authentic) problems (OECD, 2008b).

Schools and teachers should be concerned less with topic coverage and more with their learning strategy (OECD, 2008b). It is argued, therefore, that students should learn from multiple and integrated models, modules, and subjects rather than "disconnected and isolated six-week units" that are typical of the standard model (p. 49). Hattie's meta-analysis also supports focusing attention on the learning strategy by way of integrated curriculum. The author's work found significant effect sizes for integrated curriculum in elementary schools (d = 0.56) and middle schools (d = 0.57) with lesser impact on high school students (d = 0.27) (Hattie, 2009). Hattie also found that integrated curriculum has a greater effect on lower compared to middle or higher achieving students and for ethnically diverse students, when taught/implemented by more experienced teachers. Integrated curriculum is also supported (indirectly) by an empirical analysis of data collected through the Trends in International Mathematics and Science Study (TIMMS) in the 1990s. In this study, it was found that the United States, which has much more science and math content in distinct units than the other 50 countries in the survey, actually had weaker achievement compared to other countries. The authors concluded that students studying in systems with a focused and coherent curriculum will fare better than students in systems learning from crowded and distinct units (OECD, 2008b).

The expected outcome of these approaches to curriculum is improved retention of learned materials and their application in real world social and practice settings.

In sum

The use of direct and student centered instruction, summative and formative assessment, and a common, authentic and integrated curriculum with clear, challenging goals and no dead-end courses should be promoted at the school and classroom levels. These are straightforward reforms, though not without their challenges, as described above.

The instructional practices discussed here are not innovations in education; they are practices decades old (at least). The research encourages teachers to develop a large repertoire of approaches and to carefully plan how to use them. For example, direct learning approaches are most useful when teaching students basic knowledge and skills (i.e., the teacher presents factual information and actively clarifies student conceptions and misconceptions). Once students have developed an adequate knowledge base, the teacher can then use student-oriented approaches that are designed to encourage students to generate solutions to their own questions. Favouring one approach, say student-oriented practices, over another runs the risk of students engaging in theory or task-based problem solving without acquiring the basic foundational skills to organize themselves or their thinking (Klinger & Edwards, 2006; Rowe, 2007). Likewise, limiting students to directed instruction runs the risk of holding back the development of students' ability to problem solve independently, creatively, and collaboratively with their peers (OECD, 2008b).

Assessment is probably the most important factor in advancing or hindering equity in the classroom. Achievement gains associated with formative assessment "are among the largest ever reported for educational interventions" and have been found to improve equity in student outcomes by achieving general academic gains for all students (OECD, 2005, p. 2). To be clear, both summative and formative approaches have value and both are considered integral in the learning process. Success in the classroom requires that both summative and formative assessment be used in concert and practised systematically. In addition, schools and school systems that are serious about reducing failure should focus on advancing student knowledge and building confidence through useful, systematic and detailed assessment in contrast to assessment summed as a single letter (i.e., A, B, C, D, or F) or number (i.e. 9, 10). The letter or number does not provide a student much in the way of guidance on how to improve for the next assessment. The obvious challenge with this change is that education systems and education stakeholders (teachers, administrators, parents and community members) across OECD have relied on marks (or grades) for generations as part of student assessment (Black & Wiliam, 2006; OECD, 2005).

Curriculum is also recognized as important but received the least amount of attention across the literatures (compared to instruction and assessment). This could be the result of the search terms used but it may also indicate that curriculum plays less of an important role in reducing school failure than instruction and assessment. Alternatively, there may simply be fewer empirical studies available on the topic because of the complexity involved in measuring and isolating the impact of curriculum on reducing failure in the context of other variables – from pedagogy, peers, and learning environment – simultaneously impacting students in classrooms (OECD, 2008b). The critical piece is that teachers should give more attention to the development of an integrated learning strategy rather than making sure each topic is covered (again, this is something many teachers already do).

The findings in this section are drawn from a range of empirical studies, including The King's-Medway-Oxford shire Formative Assessment Project (KMOFAP) and rigorous reviews and analysis of the best available empirical research on pedagogy by the Institute of Education Sciences in the United States. Overall, the researchers were transparent, systematic and rigorous in their approach but the majority of studies reviewed employed research designs that can identify correlation and not causal relationships between variables.

The most important finding from this section of the literature review is **the need to systematically link (or integrate) instruction, assessment and curriculum**. This will not happen without challenges. For example, the more teachers integrate, differentiate and individualize their learning programs, the more difficult it will be to develop standardized assessment tools and the ability to measure and compare progress across students is lost (OECD, 2008b). Furthermore, it seems research has not yet reached a point that, by way of example, assessment strategies can be matched to specific teaching repertoires and curriculum goals in a way that is most effective for students with specific learning needs (OECD, 2005; 2008c). Though gains can still be achieved without this knowledge, system wide gains will be difficult without knowledge on how-to systematically link and intervene.

2.2 Leadership

Conventional practice across OECD countries is to head schools with a single individual referred to as either the principal, head teacher, director or other related term. The principal of a school (elementary or secondary) assumes responsibility for managing instructional quality and day-to-day operations (how much of one over the other varies across OECD countries). If school principals can take-up leadership policies and practices that improve either their instructional or administrative responsibilities, the result could be a better managed school environment with conditions set for improved student learning. This contingent nature of leadership impact on student learning is unpacked below.

How can school leaders reduce school failure? Studies investigating how leadership influences student learning point to the 'basics' or 'fundamentals' of an effective school leader: 1) sets direction that establishes a vision for the school and aligns with state and district priorities, 2) sets goals for the school that are clear, challenging, achievable, tied to district and state level goals, focused on improved student achievement and effectively communicated to staff and the community, 3) aligns school resources to set goals, 4) develops capacity in the school, such as professional learning tailored for teacher needs, and 5) protects teaching and learning time from outside distractions (Hattie, 2009; Herman et al., 2008; Leithwood et al., 2004; Seashore et al., 2010).

In terms of approach to leadership, the sources reviewed recommend that principals' take a collaborative and shared approach to decision making with staff, students, the community and other school administrators so no one individual is responsible for all requirements (Hamilton et al., 2009; OECD, 2008; Seashore et al., 2010). This approach to leadership has many names, i.e., *Collective*, *Distributed*, *Participative*. These names, in addition to the other types or brands of leadership, including *Instructional*, *Transformational*, *Moral*, *Strategic*, etc. can lead to some confusion and questioning of which type of leadership is the best approach. Leithwood et al., (2004) are sceptical of these and suggest not being confused by "leadership adjectives" since they all describe different styles of approaches of achieving the same fundamentals: "setting defensible directions and influencing members to move in those directions" (p. 6).

Aside from approaching leadership in a shared and collaborative manner, having a clear rationale supporting each action and plan, being clear and transparent about this rationale, and having access to the necessary resources to realize plans also matters.

In sum

The research strongly supports that quality leadership can play a significant role in reducing school failure. That said, the evidence related to leadership and its effects on failure and equity can be confusing to interpret. The literature on leadership is large and with so many different types identified it is difficult to know for sure how leadership models differ. This would require a complete review of the literature on leadership and its relationship to failure, which is beyond the scope of this study.

This overview of the literature suggests that while leadership influences student learning indirectly, how it does so is unclear. For example, school leaders need to assist their staff in promoting certain school and classroom conditions but the specifics of which conditions to focus on (i.e., school and class size, instructional approach) and how remains unknown. A second example, school leadership practices remain contingent on context (geographic location, level of schooling, school and district size, etc.) in a way that is not fully understood, which limits the generalizability of "effective" leadership practices (Leithwood, et al., 2004). These and other contingencies limit the ability of school districts to systematically improve conditions through leadership practice.

The review did uncover some helpful directions. School leaders must use the power in their role, as much as possible, to keep the attention, resources and the operation of the school focused on effective teaching and improving student achievement. Moreover, if school leadership decisions are made collaboratively and with strong supporting rationales, then the resulting leadership conditions in the school and classrooms should help support reduced school failure.

2.3 School and classroom learning supports

Data use in schools and classrooms, teacher professional learning, common planning time for teachers, and personalized interventions were all identified as having a supportive role in reducing school

failure. Because of their supporting role, these four topics have been organized under the thematic heading of school and classroom learning supports. The research suggests that if schools establish a *constellation of supports* that support classroom learning, then teachers and students can stay informed and focused on learning without being overburdened or having to work in isolation.

2.3.1 Data

Beginning with data use in schools and classrooms, there is a growing body of research supporting data use to inform teacher practices and day-to-day decision making in schools. For the purposes of this review, *data* is the result or output from a formal process of measurement or inquiry.

At the school level, principals could use data to inform strategic and day-to-day decision making. Examples of data relevant to the school context could be the results of state-level accountability tests, final grades from an end of term test, data on school and class size, attendance, school climate and information from surveys collected about students' well-being (Herman et al., 2008). Applications of these and other sources of data could involve the principal targeting professional learning, setting goals for the school, and monitoring the instructional practices of teachers, to name a few.

At the classroom level, teachers can use data to inform pedagogical decision making. Examples of data relevant for the classroom context include individual student achievement data (formative assessment data, state tests, and report card data), data on instructional time, disciplinary referrals and absenteeism. There are many possible applications for these kinds of data. One scenario could involve the teacher using report card data from the previous school period when developing an annual learning plan (i.e., yearlong plan outlining topics of instruction, assessment points and connections to curriculum) for the class based on the knowledge and skills of students entering the grade. In another, the teacher could use data collected through formative assessment to set individualized learning goals for students and/or select the most appropriate instructional approach.

To make effective use of data, teachers and principals will need access to "quality" data (reliable, valid and timely) and professional learning to build capacity on how-to systematically collect, prepare, and interpret school and classroom data and develop hypotheses, test approaches and modify curriculum, instruction and/or assessment accordingly (Hamilton et al., 2009). In addition, to use data effectively, schools require a strong data culture. To establish a strong data culture, school leaders should establish data teams composed of a senior school administrator, two to three teachers, one or two classroom support professionals and a qualified district level staff member who works in research, evaluation or assessment. The data team should be led by a qualified data coach and, together, the team sets the "tone" for data use in the school (Hamilton et al., 2009; Lachat & Smith, 2005). Once a data team has been established, a high priority for the team should be defining common terms used when discussing data, including 'data', 'information', 'evidence', 'achievement', and 'progress'. With a shared understanding of key terms, a plan can then be developed that articulates activities, roles and responsibilities, and sets attainable, measurable and relevant goals for the school (Hamilton et al., 2009; Lachat & Smith, 2005).

It should be noted that schools are not able to generate all the data required for effective decision-making so it is important that schools be connected with their district or regional authorities, who (hopefully) can share high-quality data (Hamilton et al., 2009; Lachat & Smith, 2005).

2.3.2 Teacher Professional learning

Teacher professional learning is an important component in any reform effort to reduce school-failure. The logic follows that by improving teacher knowledge, skills and experiences (particularly in

areas related to disadvantaged students) the result will be improved teacher performance and overall improved student achievement.

Professional learning should be well-planned, formal, systematic, deliberate and as much as possible topics should have an explicit focus on teaching specific skills. Examples of effective approaches to professional learning include observing actual classroom methods, microteaching, and video/audio feedback (Hattie, 2009; Seashore et al., 2010). Less formal and systematic approaches were found to have lesser effects, including games/simulations, coaching, printed or instructional materials and guided field trips (Hattie, 2009). Whenever possible, professional learning should bring in relevant outside experts, challenge prevailing assumptions about teaching and learning and should be supported by funded release time (Black & Wiliam, 2006; Darling-Hammond, 2010, p. 261; Gardner, 2006; Hattie, 2009; OECD, 2009). Schools also need flexibility to adapt professional learning time with school goals and needs (Hattie, 2009; Herman et al., 2008; Sebba, 2006; Wobmann, 2008).

2.3.3 Common planning time

Common planning time is also given attention for reducing failure. How does common planning time play a role? Without it, the policies and practices described above will likely remain on paper as teachers will miss the opportunity to work together. Teachers need to meet and discuss critical issues, i.e., how to use data to target instructional improvement, reach consensus on what constitutes progress in the curricula, etc. There was no mention in the literature of how often or how long common planning sessions must be, or what time in the day the planning should occur but sessions should be organized frequently and structured deeply (depth of content) enough to properly share and discuss ideas and take appropriate actions (Hamilton et al., 2009; Hattie, 2009; Herman et al., 2008; Pashler et al., 2007).

2.3.4 Personalized interventions

Incorporating personalized interventions into schools and classrooms can also serve to reduce school failure. If personalized interventions can be introduced that build student resiliency, then attention is moved away from the student's deficits towards building student confidence and ability. This shift in focus is associated with improved student engagement and reduced levels of student dropout.

Scholars report that most students that do not want to go to school also fail at school, dislike school and also dislike their teacher(s) (Hattie, 2009; Hoad et al., 2007; Levin, 2008). It stands to reason that an intervention aimed at improving the teacher-student relationship could, indirectly, reduce school failure. Hattie's meta-analysis found notable effect sizes for five teacher-student variables that would support this hypothesis. These variables are: non-directivity (d = 0.72), empathy (d = 0.67), warmth (d = 0.68), encouragement of higher order thinking (d = 0.60), and encouraging learning (d = 0.40) (Hattie, 2009, p. 118-119). So how might these teacher-student relationship variables translate into a practical classroom application? Levin (2008) talks about the positive impact on students in providing 20 minutes of concentrated time. This time can take the form of two minutes a day for 10 consecutive days or a single 20 minute conversation. The purpose of these conversations is primarily to establish a more personalized relationship between the teacher and student and this can be achieved through informal discussion focused on something the student is interested in (Levin, 2008).

In sum

Policy and practice attention was given to school and classroom learning supports as a way of reducing school failure. Building real capacity around data-use in schools and classrooms, using professional learning effectively and strategically to meet school goals, the provision of common planning

time for teachers to develop their pedagogy collaboratively, and the implementation of personalized interventions are all examples of approaches that, if used effectively, can support reduced school failure.

The link between data use in schools and classrooms is suggestive and not conclusive. There is a presumption in some articles that teachers have ready access to quality student achievement data. In reality, access to data is not guaranteed and is only one of the many challenges to data-use in schools and classrooms. Other challenges include teacher resistance toward data use, the possibility using data and negatively impacting students, data saturation and low teacher capacity to use data effectively (Herman et al., 2008; Lachat & Smith, 2005). Despite these challenges, there remains considerable optimism that data use can be used to reduce school failure. Establishing a data friendly culture and building capacity around data use are good first steps in overcoming these challenges. Part of this work could be facilitated through a data-brokering unit at the district level. An example of such a support service unit is the Evidence-Based Education and Services Team (E-BEST) of the Hamilton Wentworth District School Board in Ontario, Canada. E-BEST works with schools, departments, administrators and research partners to "coordinate and facilitate research activities, enhance capacity for evidence-based practice at all levels of the district, and model high standards for research excellence, communication, and accountability" (HWDSB, 2011). The work is intended to build the confidence and the capacity of district stakeholders to use evidence to inform daily practice and foster a positive data culture.

We know that which teaching a child receives matters. This means that equipping all teachers with appropriate knowledge, skills and experiences is essential (Hattie, 2009). We also know that without ongoing professional learning teachers fall back on what they know. The literature points to ongoing questions and concerns about how to run effective professional learning sessions. There is a clear message in the literature that professional learning should be of high-quality, challenging, sustained for as long as it is needed (often years), systematic in its approach, preferably developed by an outside expert, built from the understanding that teachers are independent professionals, and be aligned with school goals (Darling-Hammond, 2010; Grubb, 2007). These may seem like obvious findings when it comes to running professional development in schools but the research tells us they are not widely implemented.

Common planning time for teachers, like professional learning, is strongly recommended. And like professional learning, how common planning time is structured matters. There is a silence in the literature around how-to best structure common planning time, which is problematic because simply providing common planning time is no guarantee that time spent will lead to improved teacher practice.

In terms of personalized interventions, the link between personalization and student achievement is not direct and should not be understood that way. Rather, interventions that involve personalization should be thought of as an intervention that sets a positive learning relationship between the teacher and student so that other practices can have the desired take-up.

Overall the supporting evidence in this section is not strong. The evidence was either based on non-generalizable case studies (i.e., evidence on data use) or the findings were high-level and not detailed enough for school or classroom level planning. That said there was a consistent message across the literature that supports like these are needed to maintain focus on student learning. Questions remain however about how-to implement these supports effectively.

2.4 School and classroom conditions

Finally, attention was given to the role of school and classroom conditions in reducing school failure. If classrooms conditions are set at appropriate levels for learning (i.e., class size, classroom arrangement, temperature), then the empirical evidence tells us that it is **possible** these conditions **can** lead to reduced school failure.

Reduced school size has led to reduced school failure, as measured by improved student credit accumulation (Bloom et al., 2010; Hattie, 2009). For example, a study on the effect of small schools of choice[2] in the United States found that within the first three years of high school, small schools of choice enrolees earned almost one full credit more toward graduation, were 7.8 percent less likely to fail a core subject and 7 percent more likely to graduate than their control group counterparts – among other benefits (Bloom et al., 2010). These results are particularly encouraging first because they improved conditions and outcomes for low-income students of colour and the results are based on a randomized control study conducted at a large scale[3]. Hattie's meta-analysis also supports reduced school size as an intervention capable of improving student achievement, $d = 0.43$ (2009, p. 80).

There are, however, at least two important caveats with reduced school size. The first is finding the optimal school size appropriate for different demographic contexts. This review found that the optimal secondary school size ranges somewhere between 300-900 students (Bloom et al., 2010; Hattie, 2009) or as small as 250-300 students (Leithwood et al., 2004). This compares with the average secondary school size of 489 students across OECD countries, with the lowest average number less than 300 (Norway, Poland) and the largest average number 1,046 (Malaysia) (OECD, 2009)[4]. Exact school size matched to different demographic contexts, including affluence, race, urban/rural location, is not specified in the literature but there are noteworthy relationships. Affluence seems positively related with school size: the more affluent a schools' student cohort then the larger the optimal school size. Minority status seems inversely related to school size: the higher the proportion of minority students then the smaller the optimal school size (Hattie, 2009).

A second caveat finding is that reducing school size on its own does not seem to lead to improved student achievement (Darling-Hammond, 2010; Grubb, 2007; Hattie, 2009). A series of other supporting conditions need to be in place, such as purposefully organizing schools to facilitate closer student-to-student and teacher-to-student interactions and establishing a clear educational mission in the school (Bloom et al., 2010). The precise relationship between reduced school size and what supporting conditions need to be in place to reduce student failure, who the reforms should target, why, how and when remains unknown.

The literature on reduced class size reads similarly. Studies have found that reduced primary class size can lead to improved academic growth (as measured by achievement scores and graduation rates) and also non-educational benefits, including improved teacher morale, more familiar teacher-student relationships and safer schools. Tennessee's Student Teachers Achievement Ratio (STAR) project, a four-year experimental/random assignment study of reduced class size involving 11,600 different students, is one of many popular and well cited studies on the topic whose methods and results have been examined and endorsed by independent studies (Levin & Belfield, 2007; Pate-Bain et al., 1999). In broad terms, Project STAR concluded that students achieved higher graduation rates, higher test scores, and were more likely to pursue tertiary education. In 2008, a comprehensive review of the literature on the topic seems to confirm that reduced class size does provide a classroom environment for teachers to teach differently, which can lead to improved student learning particularly for students who come from low socio-economic, minority backgrounds or are low educational performers (Bascia et al., 2008; Leithwood et al., 2004). For reduced class size to work, class size must be reduced to at least 20 students per class with the greatest results between 13-20 (Bascia et al., 2008b; Hanushek, 2002; Leithwood et al., 2004).

[2] A small school of choice is academically non-selective and small allowing for reduce teacher load and common planning time.
[3] The MRDC study is an analysis of four annual cohorts (a total of 21, 085 students) who applied to one of 105 small schools of choice. The program was oversubscribed so a lottery system was used to place students who chose to be in a small school of choice, essentially creating two randomized groups: those who "won" the lottery and were assigned to a small school and those who lost and were assigned elsewhere (Bloom et al., 2010, p. 4).
[4] OECD countries participating in the Teaching and Learning International Survey conducted by OECD and published in a 2009 report.

Like school size, rigorous evaluation studies of reduced class size are rare and there are caveat findings (Levin & Belfield, 2007). Class size policy must remain in place for a number of years (minimum of three to four), changes in classroom practice must accompany the reform, such as ongoing professional learning (for new and veteran teachers) focused on the use of effective instructional practices, etc. (Bascia et al., 2008b; Leithwood et al., 2004). In addition, studies on the topic have mostly focused on the relationship between class size and student achievement from a quantitative perspective (i.e., student and teacher numbers) without exploring: what is actually taking place in the classroom? How (if at all) it is impacting teaching and learning? Does it improve overall student success? And if so, how (Bascia & Faubert, In Press)? A recent evaluation study examining the early impact of reduced primary class size in Ontario, Canada (where 90% of kindergarten to grade 3 classroom sizes were required to have 20 students or less) found that reduced class size improved teaching and learning conditions in primary grades; in particular, the use of classroom space and teacher pedagogical approach. When researchers compared classroom arrangements in primary versus upper grade classrooms across multiple schools sites included in the evaluation, it seemed that the reconfigured space in primary classrooms (the result of reduced class size) enabled the use of a wider range of pedagogical approaches that put the learner centre stage. This is an important implication for policy (Bascia & Faubert, In Press) because such changes have been linked to "deep learning" (OECD, 2008, p. 32). This level of observation and analysis at the classroom level remains uncommon in reduced class size studies.

This review found that reduced class size alone is likely to result in only modest gains. Hanushek (2002) concludes that any effects of overall class size reduction will be small and possibly benefit mostly those students living in high-poverty areas and students who are underperforming. Hattie's (2009) meta-analysis seems to echo Hanushek's conclusion. Specifically, the author finds that studies reporting greater class size effects were more related to improvements in teacher and student working/learning conditions, while studies with lower overall effects focused primarily on student learning. Hattie's analysis gives class size reduction an overall effect of $d = 0.21$.

The discussion on school and classroom conditions also covers physical conditions (temperature, acoustics and lighting). Research indicates that maintaining stable classroom conditions is important for classroom management and student learning. School and classroom temperature should be in the range of 19-23 degrees Celsius (67 – 74 degrees Fahrenheit) and acoustics around 40 decibels (or the hum of a refrigerator) (Earthman, 2002).

In sum

School and class size reduction has intuitive appeal to administrators, teachers, students and parents. But a closer inspection of the research points to a mixed relationship between small schools and classrooms and expectations of achieving significant positive effects (Stevenson, 2006; Wobmann, 2008). The research shows that reduced school and class size can increase the opportunities to learn by increasing the possibilities for student-student and student-teacher interactions. It is both the frequency and quality of these interactions in small schools and classrooms – individualized instruction, regular formative assessment, better use of space, different grouping strategies and activities – that matters and these interactions do not change necessarily with changes in school or classroom size. Moving from one school or class size to another requires a shift in what it means to be an effective teacher (Hattie, 2009). The literature currently does not provide answers, supported by robust evidence, to these questions.

Furthermore, reducing school or class size is costly as additional teachers, custodial and clerical staffs need to be hired, new facilities need to be built or rented, and the cost of utilities is higher (Hanushek, 2002). For example, primary class size reduction cost the education system in Ontario, Canada approximately $400 million CAD for new primary teachers and over $716 million CAD for capital improvements, totalling $1.1 billion dollars CAD (Bascia et al., 2008b). In the Bloom et al., (2010) study,

the project benefited from tens of millions of USD worth of funding from philanthropies. In short, the fiscal costs of school or class size reduction may be too high, particularly in the current environment of fiscal restraint across all OECD countries.

2.5 Summary overview of policies and practices aimed at overcoming school failure and improving equity

The table below provides a classification of the policies and practices supported by empirical research to reduce school failure. The organization of the table parallels Section 2 of this review. The summary overview begins with the thematic category, followed by what policy or practice impacts school failure, who are the actors with influence, how the policy or practice reduces failure, and why it has impact.

Thematic Category	What - policy / practice impacts school failure	Who – actors with influence	How - the policy / practice reduces school failure	Why – the policy /practice has impact
Pedagogy	Instruction	Teachers	Using a combination of direct and student-oriented approaches to instruction with "common sense" (see section 2.1 Pedagogy, In Sum)	The research points to policies and practices in assessment and instruction as having the strongest and most direct impact in reducing school failure. Instruction, curriculum and assessment are very inter-connected and it is recommended they be systematically linked for greatest impact.
	Assessment		Combined use of summative and formative assessment, with a greater emphasis on formative assessment.	
	Curriculum		Establish a common and integrated curriculum with courses that lead to real educational or career options; focus attention on developing a learning strategy with clear learning goals and coherent across lessons.	
Leadership	Instruction	Principals Teachers District administrators Parents Community leaders	Keep the attention, resources and operation of the school focused on effective teaching and improved student achievement. Principal should lead instructional improvement, make strategic use of professional learning, and promote a strong data culture in the school.	The research strongly supports that quality leadership can play a significant role in reducing school failure. However, the influence of leadership on student learning (though important) is indirect and there are many contingencies in the research findings.
	Administration		The principal should develop a vision for the school and communicate this plan to teachers, parents and community members. The principal should align school goals with district and state level policies; resourcing decisions made at the school level should also be aligned with school goals. Leadership decisions should be made collaboratively (teachers, students, community, and other school administrators) and each action should be supported by a strong rationale.	

Thematic Category	What - policy / practice impacts school failure	Who – actors with influence	How - the policy / practice reduces school failure	Why – the policy /practice has impact
School and Classroom Learning Supports	Data use in schools and classrooms	Teachers Principals District Administrators Students	Build effective teacher, principal and student capacity around data use and foster a positive data culture.	Establishing supports for teaching and learning in the classroom enables capacity and collaboration required to focus on school goals in an environment of diverse and multiple challenges.
	Professional learning		Run effective professional learning sessions that are strategically focused on school goals.	
	Common planning time		Make it possible for teachers to develop pedagogy and problem solve collaboratively during school hours.	
	School based interventions		Implement personalized interventions that build student resiliency and confidence.	
School and Classroom Conditions	Reduced school size	Teachers	School size is limited to 300-900 for secondary schools, smaller for elementary schools. The more affluent a schools' student cohort then the larger the optimal school size. The higher the proportion of minority students then the smaller the optimal school size. Changes in school size should be accompanied by reforms in the school and classroom.	It is both the frequency and quality of student/student and student/teacher interactions in small schools and classrooms that matters and these interactions do not change necessarily with changes in school or classroom size. Moving from one size to another requires a shift in what it means to be an effective teacher.
	Reduced class size		The research shows that reduced school and class size can increase the opportunities to learn by increasing the possibilities for student/student and student/teacher interactions. Class sizes of 20 students or less must be in place for a minimum of three to four years and changes in classroom practice must accompany the reform.	
	Physical conditions		Maintaining appropriate physical classroom conditions is important for classroom management and student learning.	

3. Challenges and considerations in the design and implementation

In section 2, the strength of the supporting empirical evidence, considerations and challenges specific to the design and implementation of policies and practices were discussed. In this section, a broader discussion of the challenges and considerations for design and implementation is provided.

Resourcing

Resourcing is a major consideration in the design and implementation of policies and practices. At the most basic level of resourcing, most of the policies and practices discussed here require start-up funds to hire additional staff, improve professional learning, and/or secure new facilities or equipment. And of course, once the new initiative is up and running, additional sustained funding is often required: most interventions must be in place for several years for results, like change in teacher practice, to be realized.

Human capital is a second, more complex resourcing consideration since qualified teachers, instructional support personnel, and administrators (trained in evidence based pedagogies and leadership practices) are always in short supply (Darling-Hammond & Friedlaender, 2008; OECD, 2009). More complex yet is how resourcing differences arise from critical differences in demography and political culture. For example, larger districts are generally better able to redirect both funding and human capital resources to support the implementation of changes better than smaller districts (Sebba, 2006; Bascia et al., 2008b). In terms of political culture, districts that take a punitive approach (in the form of reduced resources to schools) for lower performance make already difficult resourcing decisions even more challenging.

Adequately resourcing policies and programs can cost billions of dollars (USD) and consume large amounts of human capital; as is the case with reduced school and classroom size. To pay for these, systems must either **spend** more or **redistribute** resources from other initiatives.

Spending more on policies and programs does not automatically lead to improve outcomes (Grubb, 2009). This finding and the current state of fiscal restraint in the public sector means governments are increasingly reluctant to significantly increase spending in education because the return on investment is suspect. In terms of resource redistribution, researchers in education finance have identified several areas in which education funds are misspent – unnecessary inputs like textbooks, supplies or computers, poorly developed professional learning programs, and attempted reforms backed by insufficient resources (Grubb, 2009). Talk of redistribution, however, quickly becomes contentious as salaries (teachers, administrators) and programs (including special education) make up the bulk of spending in education. For complex reasons, using resources effectively in education does not seem to be part of the culture of many countries. There is a general lack of quality cost/benefit analyses of different educational policies and programs at the school and district level, meaning that schools and districts are making decisions with minimal attention to the efficiency or effectiveness of education outcomes (Hattie, 2009; Wobmann, 2009).

Issues concerning the transferability of policies and practices across jurisdictions to another

There are at least three major issues with the assumption of transferability of a policy or practice from one jurisdiction to another: 1) the effects of cultural differences across countries, 2) the effects of variability in the structure of governance in education, and 3) the effects of variability in available resources across and within jurisdictions.

The assumption that educational innovations in one context can be transferred over into another (whether the transfer be from one country, district or individual) is not unproblematic. One such problem is posed by cultural differences, which present particular challenges in the interpretation, selection, take-up,

implementation, and sustainability of policy and practice. The cultural differences within and between OECD countries are real and considerable. Stakeholders from each OECD country will interpret the research presented in this review with their own understanding of the critical challenges contributing to school failure and factors undermining equity in their country. Based on their interpretation, stakeholders will select policies and practices viewed as appropriate to the social context (i.e., culture, values) of their respective country/educational jurisdiction. That culture affects educational decision making is a point well documented in the literature reviewed. For example, Edwards et al. (2007) found that educational policy makers and decision makers from more individualistic societies tend to overemphasize the individual's ability to overcome failure and underestimate the relationship between the role of the school or classroom and their impact on student outcomes. The result is an interpretation and selection of policies and practices that reflect the individualistic culture and values of society but may not address the causes of failure if the causes are institutional or systemic. A second example, Levin (2008b) talks about the Ontario government's experience shifting education culture in the province from one of antagonism to respect for teachers as key to ensuring the take-up, implementation and sustainability of reforms aimed at building system wide capacity and steering focus on student outcomes. The lesson here is system wide reform should begin by establishing or reinforcing a culture of trust and partnership or reform efforts will likely be hampered.

Cultural differences in teacher beliefs also impact the take-up, implementation, and sustainability of policies and practices empirically established as effective. Across OECD countries, teacher beliefs about issues, like instruction, are dependent on cultural context and traditions. These beliefs are formed early on in teacher formation and typically remain stable over time (OECD, 2009). This suggests that even the findings from the most rigorous experimental studies may not be transferable from one state to another, yet alone a country, not because the research findings are wrong but rather because the effective practice may not be culturally appropriate in another country. Such a case was reported by Black & Wiliam who found that the successful implementation and take-up of formative assessment was heavily dependent on the social and educational cultures of different jurisdictions (Black & Wiliam, 2006). In her recent book on America's commitment to quality and equity in education, Darling-Hammond (2010) remarked that "replication efforts [presumably in the United States] have an inglorious history, largely because they quickly run up against differences in staff knowledge and...contexts of receiving schools" (p. 272) – which are essentially differences in school and classroom cultures. She notes that in worst-case scenarios, "purportedly effective practices" do not work immediately and teachers revert back to old approaches to teaching because these are familiar and the prevailing teaching culture supports the practice. These examples highlight the need for caution. If cultural differences/preferences are either dismissed or given inadequate attention in considerations of transferability then, despite the strongest possible evidence, the take-up, implementation, and sustainability of any intervention will be undermined and likely fail. Appreciating cultural differences (language, goals, purpose of education) and addressing possible tensions when introducing any reform are important steps in fostering a culture open to change (OECD, 2005; 2008c).

A second problem is the allocation of power over governance of education, and, relatedly, the distribution of roles and responsibilities across levels of government in a jurisdiction. There is great variability in the structure of governance of education across OECD countries. A related issue is the role of the private sector in education. There is considerable variability in the nature and degree of private sector involvement in, for example, the operation and management of schools and many permutations of public sector-private sector collaboration and partnership. Whether the structure of governance of education and the role the private sector plays allows for flexibility around control of decision-making, policy formation, and resource expenditure will significantly affect whether a policy or practice can be successfully transferred. While clearly a significant issue, this review of the literature suggests that it generally not given adequate attention in considering transfer of policy and practice.

This next point is closely related to the earlier discussion on resourcing. Much like resourcing receives too little consideration in design and implementation, the variability in resources across and within jurisdiction receives too little attention when discussing the transferability policies and practices. Previously mentioned qualified human capital and funding are always in short supply but communities also vary in terms of access to transit and other social services, as well community resources including local mentors and businesses, particularly when one compares urban and rural communities. The literature is clear that these and other resources significantly affect the successful implementation, take-up and institutionalization of policies and practices. By way of example, implementing innovative programs, like the Dual Credits program[5] in Ontario, Canada, relies on the availability of all these resources for program success. Schools in urban, resource rich areas can take advantage of the ready availability of qualified teachers, local colleges, adult mentors, community facilities, public transit and local businesses to engage in practice based learning. Schools in rural, resource poor communities may find it difficult to attract teachers to the area; colleges may not be located within a convenient travel radius for students and public transit non-existent. Successful initiatives, therefore, cannot be assumed to work across jurisdiction or even in the same community. A proper feasibility analysis of available resources is an important step when considerability transferability, even of those programs widely recognized and supported by robust evidence to be effective.

Recognizing the challenges of policy and practice reform

There are at least two serious challenges confronting policies and practices. First, policies or practices implemented at the classroom levels are frequently designed and developed at much higher levels in education systems. The greater the distance between program development and implementation, the greater the possibility that other factors will impact and affect the outcomes. In Leithwood et al. (2004) the authors note that the already contingent relationship between leadership practices and influencing student learning becomes even more complex the larger the organization or system. For example, the chain of variables linking leadership practices with student learning is much longer for district leaders than for school leaders, which suggests the further away a leader or policy is from the classroom the more difficult it will be to achieve direct and causal influence. Many such extraneous causal factors are difficult to control.

A second challenge is the overly-ambitious expectations of what can be achieved through the reform of policy and practice at the classroom level. Professional learning and building capacity to use data in schools is a good example. Professional learning can be an effective way to refresh previously acquired knowledge or mastered skills or to introduce new knowledge and skills that builds on an already strong foundation. But the ability to scope an issue/question, define a methodology, collect and analyze data and then implement solutions is not a standard skill set across all teachers. These skills must be developed over time and with the support of trained experts who can correct misconceptions and provide appropriate direction. If teachers are expected to use data, then policies should be set requiring teachers to have data training prior to entering teacher training programs or make courses available on data use in schools and classroom a requisite component of teacher professional programs. A similar comment applies to the use of technology.

None of the challenges discussed above, individually or as a group, suggest that implementing policies and practices developed in another jurisdiction is impossible. In fact, a primary reason for examining other education systems is to find ways to achieve political, social and economic gains in the home jurisdiction (Bray, 1999). Section 4 of this review offers policy and practice connections that can be used as a beginning for effective school and classroom level reform aimed at reducing school failure.

[5] The Ontario Dual Credits Program focuses on students who have the potential to succeed but are at risk of not graduating from upper secondary school. Students participate in apprenticeship training and postsecondary courses, earning dual credits that count towards both their high school diploma and their postsecondary diploma, degree or apprenticeship certification (Quote: Ontario, 2010).

4. Connections across policies and practices

The importance of alignment

The first connection is the need for alignment across areas of policy and practice and alignment across the levels of education systems in their implementation (OECD, 2005; 2010c). The research shows that curriculum instruction, assessment, leadership and data use have the most direct impact on the core business of teaching and student learning, relative to the other policies and practices. Systematic alignment of policy and practice initiatives across these five areas in conjunctions with systematic alignment of implementation across all levels of school systems significantly improves the likelihood of success and increases impact. The consequence of not aligning and systematizing will be the uneven take-up of evidence based policies and practices in classrooms across the system which can lead to potholes along educational pathways that less-academically inclined students can succumb to. The interconnectedness across the five areas also suggests that change of any one area necessitates consideration of how the change will impact the other four. For example, changes in instructional practice will surely impact on assessment practices and data use.

The importance of a support system

A second and related connection is the need for a well-developed constellation of supports for teachers and school leaders to maintain focus on student learning in schools and classrooms. To accomplish this, supports must be in place for teachers to build capacity, problem solve, and innovate. The literature suggests that part of this support framework should include teacher professional learning, common planning time for teachers, and personalized student interventions. For example, if teachers are asked to introduce formative assessment processes in their classrooms, then high-quality and sustained professional learning must be provided to demonstrate how assessment can be systematically integrated into curriculum and instructional processes.

The importance of layered and selected activities

A third connection is to think strategically about the selection of policies and practices. One or two reforms working in isolation is unlikely to reduce school failure (OECD 2010c). Combining mutually supporting policies and practices into a web will greatly increase overall impact. The consequence of poor selection was observed in the 2010 Wallace study on leadership. The case was of a district that "picked and chose" two of the ten possible recommended practice reforms. Not only did the district not realize improved results but research participants reported a form of reform "backfire" as a result of poor strategic selection and implementation of reforms (Seashore et al., 2010). A related issue is the need to select policies and practices that are suitable for the realities of the school and classroom context. For example, if 95% of a districts K-12 class sizes range between 40-50 students, implementing formative assessment practices that involve a lot of sustained individualized feedback to students will likely result in teachers being over-worked and not able to implement the practice meaningfully. Schools should also select policies and practices that work with district and state goals while simultaneously work to reduce school failure.

The importance of early intervention

A fourth connection across policies and practices is the importance of initiatives intended to impact student achievement early on. There are two good reasons to do this. As students move on in the elementary and secondary educational formation, it becomes more difficult to find policies and practices that both reduce failure and improve equity. Students are typically tracked (i.e., by ability) by the time they reach secondary school and gaps in knowledge and skills widen and deepen as students move further along

in their education. This means the policies and practices intended to reduce failure or improve equity must impact a greater number of targets that become increasingly more complex to resolve with each passing year; and each passing year means one less year to influence student outcomes. Across OECD countries, K-12 education has near universal enrolment (at least until grade 10 or lower secondary schooling) so any policy that increases the quality of schooling in terms of students' cognitive and non-cognitive skills in these years (with greatest impact in the early years of K-3) should be considered (Wobmann, 2008). Such interventions can also produce a multiplier effect as a student's knowledge and/or skills will not only increase at the stage where/when the approach is introduced, it also better prepares the student to acquire additional skills at later stages (Grubb, 2007; Wobmann, 2008).

The importance of resources and costing

There is an almost complete absence of discussion of resource requirements and the financial costs of undertaking the policy and practice initiatives identified. This is unfortunate because the resources invested in establishing and maintaining programs is not trivial. Transparent, detailed and accurate estimates of costs are an all-too-rare situation. However, the importance of costing the resource requirements of initiatives under consideration and assessing costs against anticipated outcomes and impact is critical, particularly in the present fiscally-constrained environment.

An over-arching piece to consider relevant to all five connections is the need to think systematically at all stages along the way from policy to practice (conception, implementation, and evaluation) and at all levels of the system. Policy makers and educational administrators should consider which policy and practices work best for themselves in a systematic manner (based on sound comparative methodology) to determine if the proposed changes can be situated and configured to the context.

5. Conclusion

This literature review examined the most relevant empirical literature on the subject of in-school practices for overcoming school failure. The literature review sought to answer two general questions: 1) what are the policies and practices that best support reduced school failure and improving equity? And, 2) what is the empirical evidence of their impact on reducing school failure and improving equity?

To answer the first question, a search of the literature was conducted. The search revealed significant debate and research activity around key policies and practices empirically assessed to reduce school failure and improve equity. The review finds that pedagogical knowledge matters and is central to reducing school failure and achieving equitable outcomes. The evidence points to the great potential in systematically linking instruction and assessment strategies with curriculum. School leaders (administrators, teachers) play a critical role in fostering conditions for success by encouraging effective pedagogical knowledge, establishing effective learning supports, and maintaining appropriate learning conditions while minimizing distractions. The role of data in decision making is emerging as a critical element in advancing students to their goals, school planning, and shaping/aligning strategy, policy and practice at all levels of an education system.

The amount of empirically assessed research available on each topic varies. For example, empirical research on effective instructional and assessment practices far outweighed the empirical research on curriculum practices. High-quality professional learning is widely discussed in the literature as essential to improve student learning but there was no mention of specific professional learning topics that should be given greater priority or where and when to provide the training, who is to receive it, and how the training ought to be delivered. Overall, what emerged from the literature search were the over-arching and interconnected themes of pedagogy, leadership, school and classroom learning supports, and school and classroom conditions which also make-up the thematic categories of the summative overview in Section 2.

With respect to question two, all topics faced similar issues around research design and the claim of transferability. The majority of studies reviewed employed research designs that can identify correlation and not causal relationships between variables. In addition, the United States is the source of most empirical research on the topic, raising further questions as to the transferability of the policies and practices recommended.

Empirical research providing unambiguous answers to *what*, *when*, *why*, and *how* policies and practices are effective in improving student achievement remains limited. For example, questions such as, how should a teacher adapt their instructional approach when class size is reduced from 30 students to 15? If a teacher employs a grouping approach to student learning, what are the specific strategies that must be taught to students by the teacher before effective group learning can take place? If reducing the size of a school or classroom alone is not guaranteed to be effective in improving student outcomes, then what other reforms must follow to help create the conditions for improvement? There is little empirical evidence to answer many critical questions like these that, ironically, are at the core of reducing school failure.

The research is clear that there is no silver bullet that can reduce school failure and improve equity. Instead, a layered approach of policies and practices at the school and classroom level is recommended, some intended to focus on the core of teaching and learning (instruction, curriculum, assessment, leadership and data use) while others create the network of supports required to maintain teacher focus on student learning (professional learning, reduced school and classroom size, and interventions that build resiliency). Crucially, these layers must form part of a larger design properly aligned to support efforts across all levels of the education system.

REFERENCES

Bascia, N. and E. Fredua-Kwarteng (2008), "Class Size Reduction: What the literature suggests about what works", *Canadian Education Association*.

Bascia, N., C. Connelly, J. Flessa, B. Mascall, P. Chindalo, B. Faubert, E. Fredua-Kwarteng, J. Leung and C. Rottman (2008b), "Ontario's Primary Class Size Reduction Initiative: Report on Early Implementation", *Canadian Education Association*.

Bascia, N. and B. Faubert (in press), "Physical space, spatiality, and policy space: How class size reduction affects teaching and learning", *Leadership and Policy In schools.*

Black, P. and D. Wiliam (2005), "Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices," *The Curriculum Journal*, vol. 16, No. 2, pp. 249-261.

Black, P. and D. Wiliam (2006), "Assessment for learning in the classroom", in J. Gardner (ed.), *Assessment and Learning*, Sage Publications, London, pp. 9-25.

Bloom, H., S. Thompson and R. Unterman (2010), "Transforming the High School Experience – How New York City's New Small Schools Are Boosting Student Achievement and Graduation Rates: Executive Summary", mdrc Building Knowledge to Improve Social Policy. http://www.mdrc.org/publications/560/execsum.pdf

Bray, M. (1999), "Methodology and focus in comparative education", in M. Bray and R. Koo (eds.), *Education and society in Hong Kong and Macau: comparative perspectives on continuity and change*, Comparative Education Research Centre, University of Hong Kong, pp. 209–228.

Darling-Hammond, L. and D. Friedlaender (2008), "Creating excellent and Equitable Schools: The right design features and policies can promote exceptional high schools on a broad scale", *Educational Leadership*, Vol. 65, No. 8.

Darling-Hammond, L. (2010). "Organizing for Success: From Inequality to Quality. The flat world and education: How America's commitment to equity will determine our future", Teachers College Press, New York, pp. 234-277.

Earl, L. (2003), "Assessment as Learning: Using Classroom Assessment to Maximize Student Learning" T. Guskey and R. Marzano (eds.), Corwin Press Inc., *pp. 1-152*.

Earthman, G. (2002), "School facility conditions and student academic achievement", E-scholarship repository, http://repositories.cdlib.org/idea/wws/wws-rr008-1002

Edwards, O., V. Mumford, M. Shillingford, and R. Serra-Roldan, (2007), "Development assets: A prevention framework for students considered at risk", *Children & Schools*, Vol. 29, N. 3.

Gardner, J. (2006), "Assessment and learning: An introduction", in J. Gardner (ed.), *Assessment and Learning*, Sage Publications, London, pp. 1-5.

Grubb, N. (2007), "Dynamic inequality and intervention: Lessons from a small country," *Phi Delta Kappan*, Vol. 89, No. 2, pp. 105-114.

Hamilton, L., R. Halverson, S. Jackson, E. Mandinach, J. Supovitz, and J. Wayman (2009), *Using Student Achievement Data to Support Instructional Decision Making*, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, Washington, DC, http://uies.ed.gov/ncee/wwc/publications/practiceguides/

Hamilton Wentworth District School Board (2011). E-BEST @ HWDSB, http://www.hwdsb.on.ca/e-best/index.aspx

Hanushek, E.A. (2002), "Evidence, politics, and the class size debate", in L. Mishel and R. Rothstein (eds.), *The Class Size Debate*, Economic Policy Institute, Washington DC, pp. 37-65.

Harlen, W. (2006), "On the relationship between assessment for formative and summative purposes", in John Gardner (Ed.), *Assessment and Learning*, Sage Publications, London, pp. 103-117.

Hattie, J. (2009), *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*, Routledge, United Kingdom.

Herman, R., P. Dawson, T. Dee, J. Greene, R. Maynard, S. Redding, and M. Darwin (2008), "Turning Around Chronically Low-Performing Schools: A practice guide", National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, DC, http://ies.ed.gov/ncee/wwc/publications/practiceguides

Hoad, K., J. Munro, C. Pearn, N. Purdie, L. Ellis, and K. Rowe (2007), "Working Out What Works Training and Resource Manual: A teacher professional development program designed to support teachers to improve literacy and numeracy outcomes for students with learning difficulties in Years 4, 5 and 6 (2nd edition)", published by Australian Council for Educational Research, Australia, http://www.dest.gov.au/NR/rdonlyres/4781A2FC-A2B4-4C65-88D0-30C5507E4ADE/18998/WOWW2ndEdition.pdf

Klinger, J. and P. Edwards (2006), "Cultural considerations with response to intervention models", *Reading Research Quarterly*, Vol. 41, No. 1, pp. 108-117.

Lachat, M. and S. Smith (2005), "Practices that support data use in urban high-schools", *Journal of Education for Students Placed at Risk*, Vol. 10, No. 3, pp. 333-349.

Leithwood, K., K. S. Louis, S. Anderson, and K. Wahlstrom (2004), "Review of research: How leadership influences student learning," Wallace Foundation, http://www.wallacefoundation.org/SiteCollectionDocuments/WF/Knowledge%20Center/Attachments/PDF/ReviewofResearch-LearningFromLeadership.pdf

Levin, B. (2003), "Approaches to equity in policy for lifelong learning: A paper for the OECD", Commissioned by the Organization for Economic Cooperation and Development, Paris, France.

Levin, B. (2008), "20 minutes to change a life?", *Phi Delta Kappan*, Vol. 90, No. 5, pp. 384-385.

Levin, B. (ed.) (2008b), *How to change 5000 schools: A practical and positive approach for leading change at every level*, Harvard Education Press, Cambridge.

Levin, H. and C. Belfield (2007), "Educational interventions to raise high school graduation rates," in C. Belfied and H. Levin (Eds.), *The Price We Pay: Economic and social consequences of inadequate education*, Brookings Institution Press, Washington DC, pp. 177-199.

OECD (1998), *Overcoming Failure at School*, Published by the Organization for Economic Cooperation and Development, Paris, France.

OECD (2005), "Formative assessment: Improving learning in secondary classrooms Policy Brief*"*, *OECD Observer*, http://www.oecd.org/dataoecd/19/31/35661078.pdf

OECD (2007), "No More Failures: Ten Steps to Equity in Education, Summary Policy and Recommendations", Published by Organization for Economic Cooperation and Development, Paris, France, http://www.oecd.org/dataoecd/50/18/38692453.pdf

OECD (2008), "Improving School Leadership, Volume 1: Policy and Practice," Published by Organization for Economic Cooperation and Development, Paris, France.

OECD (2008b), "Innovating to Learn, Learning to Innovate", *OECD Publishing*, http://www.oecd.org/document/7/0,3343,en_2649_35845581_41656455_1_1_1_37455,00.html doi: 10.1787/9789264047983-en

OECD (2008c), "Assessment for learning: Formative assessment," OECD/CERI International Conference Learning in the 21st century: Research, Innovation and Policy, the Centre for Education Research and Innovation, Paris, France, http://www.oecd.org/dataoecd/19/31/40600533.pdf

OECD (2009), "Creating effective teaching and learning environments: First results from Teaching and Learning International Survey," the Organization for Economic Cooperation and Development, Paris, France, http://www.oecd.org/dataoecd/17/51/43023606.pdf

OECD (2010), "Overcoming School Failure: Policies that Work: OECD Project Description", the Organization for Economic Development, Paris, France.

OECD (2010b), "Education at a Glance 2010: OECD Indicators," the Organization for Economic Cooperation and Development, Paris, France, http://www.oecd.org/dataoecd/45/39/45926093.pdf

OECD (2010c), "Improving schools: strategies for action in Mexico," the Organization for Economic Cooperation and Development, Paris, France, http://www.oecd-ilibrary.org/education/improving-schools_9789264087040-en

Ontario Ministry of Education (2010), "Dual Credits Program", Queen's Printer for Ontario, 2010, http://www.edu.gov.on.ca/eng/teachers/studentsuccess/DualCredit_FactSheet.pdf

Pashler, H., P. Bain, B. Bottge, A. Graesser, K. Koedinger, M. McDaniel, and J. Metcalfe (2007), "Organizing Instruction and Study to Improve Student Learning", National Center for Education Research, Institute of Education Sciences, U.S. Department of Education, Washington, DC, http://ncer.ed.gov

Pate-Bain, H., B. DeWayne Fulton, and J. Boyd-Zaharias (1999), "Effects of class size reduction in the early grades (K-3) on high school performance: Preliminary results (1999) from Project STAR,

Tennessee's longitudinal class-size study", Health and Education Research Operative Services (HEROS), http://www.heros-inc.org/star-hs-p.pdf

Pullin, D. (2008), "Assessment, Equity and Opportunity to Learn", in P. Moss, D. Pullin, J. Paul Gee, E. Haertel, and L. J. Young (eds.), *Assessment, Equity and Opportunity to Learn*, Cambridge University Press, New York, pp. 136-169.

Rowe, K. (2007), "The imperative of evidence-based instructional leadership: Building capacity within professional learning communities via a focus on effective teaching practice", Background paper to keynote address presented at the Sixth International Conference on Educational Leadership, University of Wollongong, 15-16 February 2007.

Seashore-Louis, K., K. Leithwood, L. Wahlstrom, and S. Anderson (2010), "Learning from Leadership Project – Investigating the links to improved student learning. Final Report of Research Findings," Commissioned by The Wallace Foundation, http://www.wallacefoundation.org/KnowledgeCenter/KnowledgeTopics/CurrentAreasofFocus/EducationLeadership/Documents/Learning-from-Leadership-Investigating-Links-Final-Report.pdf

Sebba, J. (2006), "Policy and practice in assessment for learning: the Experience of selected OECD countries", in J. Gardner (ed.), *Assessment and Learning*, Sage Publications, London, pp. 333-351).

Stevenson, K. (2006). *Major educational trends shaping school planning and design: 2007.*

Waxman, H., Y. Padron, and K. Arnold (2001), "Effective instructional practices for students placed at risk of academic failure," in G. Borman, S. Stringfield, and R. Slavin (eds.), *Title I Compensatory Education at the CrossRoads*, Lawrence Erlbaum Associates Publishers, New Jersey, pp. 137-170.

Wiliam, D. (2010),"What counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment", *Review of Research in Education*, Vol. 34.

Wobmann, L. (2008), "Efficiency and equity in European education and training policies," *International Tax Public Finance*, Vol. 15, pp. 99-230.

ANNEX A: RESEARCH QUESTIONS AND TERMS OF REFERENCE

The age/grade educational model of student progression is at least a century old. Variants of this model exist across the world but essentialy a group of students is placed in one classroom, with one instructor whose role is to transmit knowledge and develop skills. There are two basic outcomes for students in this model: if a student successfully meets the achievement requirements of the grade or course level (criteria typically determined by the classroom teacher based on his/her interpretation of the curricula), the student will be permitted to advance to the next grade or course level with their age cohort. If the student does not meet the requirements, she or he will be failed – which often means repeating a grade or course level if the student remains in the school system.

The theory of and belief in equality in education has been around since the establishment of the age/grade progression model. Education in this view is a public good. To meet the basic principles of equality in education, the state had to make (at least) primary schooling accessible to a student at no cost to the individual. Equality in education has evolved since it was first conceived with the expansion of civil rights (positive rights) to include all students regardless of gender or race. The principle of equality in education and the age-grade progression model seemed to work well together. The state met its requirement to provide all students with free access to a school and, thereby an "education". The outcome was up to the individual: if a student made it through to the end, she or he had succeeded; if not, she or he had failed. In principle, all students were provided equal opportunity to education, regardless of their socio-economic status, gender, ethnicity, etc. For decades, millions of students advanced through the grade levels, while millions of other students failed to progress. These outcomes were widely regarded by school officials, policy makers and the public as a natural and desirable consequence of schools and teachers doing their job.

Through the 1950s the post-war expansion of the social-welfare state was accompanied by massive economic investment in public services, including education. Human capital theory posited a link between education and improved economic value for the state. There was a growing awareness amongst policymakers, scholars and the public that failure in schools might not be in the economic interests of the country. The civil rights and student movements of the 1960s challenged the long held belief that student failure was exclusively a deficiency on the part of the student in their intellectual performance, behaviour or effort (or a combination therefore). The social consensus on the age/grade educational model began to erode as the relationship between failure and equity (or inequity) in education became a growing concern.

A new understanding of failure emerged in scholarly research in the 1960s. Over the decades, researchers investigating failure at school (i.e., low-achievement, not attaining minimum standard of performance) found a positive relationship between student dropout rates and challenges in adulthood (i.e., unemployment, lower civic participation) as well as high public financial costs (for health, justice and social welfare). In addition, scholars identified patterns in failure at school, with specific groups of students (i.e., racialized minorities, special needs, and lower socio-economic status) being over-represented. As a result, policymakers began paying attention to a wider range of contributing factors in the 1960s, as a growing body of research pointed to the role of teachers, classrooms, schools, districts and the state all contributing to the failure of students in schools.

The decades of the 1980s, 1990s, and 2000s were critical for research and policy regarding failure in education. Researchers in economics, psychology, sociology, political science and education began

deepening their investigating into the results of failure and why and how students fail. In addition, the emergence and consolidation of the learning sciences (cognitive science, educational psychology, computer science, anthropology, sociology, education studies, etc.) and evidence regarding how individuals learn pointed to the strong influence of peers, teachers, school environment and other context variables on learning (as opposed to only individual effort and merit). Together these pointed to the need for significant reforms to all levels of the education system in an effort to raise overall standards of educational performance, reduce gaps in student achievement outcomes and return public confidence to publicly funded education. Now more than ever, the focus in both policy and academic circles is the link between failure and equity and the push is on to improve equity in schools to reduce school failure. The obvious challenge of meeting this mandate is having to reform an institutional and organizational model established in the mid 19th century when the purpose of education was to sort students into different paths as quickly, efficiently and cheaply as possible.

The purpose of this report is to review the body of literature concerned with reducing school failure by improving equity in schools and classrooms. The literature review will be used to inform the OECD Project *Overcoming School Failure: Policies that Work* and hopefully, future educational reforms across OECD countries that aim to improve equity in schools and to reduce failure. Specifically, the terms of reference for this assignment set the focus of the literature review on the most relevant literature on the subject of in-school policies and practices for overcoming school failure and answering two general questions: 1) what are the policies and practices that best support reducing school failure and improving equity? And 2) what is the empirical evidence of their impact on reducing school failure and improving equity?

ANNEX B: SEARCH METHDOLOGY

Step one was to delineate/determine the search terms. Inclusion and exclusion criteria set for the literature review required that the review include empirical policies and practices that reduce school failure and improve equity focusing on classrooms and schools and to lesser extent school districts as units of analysis (operational/organizational levels). By the terms of reference of this assignment, topics excluded from this review include system-level policies and specific policies and practices on special education learning needs, English language learners or other specific student groups. In addition, empirical studies on policies and practices concerning chartered schools, religious schools, private schools, bilingual schools were omitted. Finally, specific empirical studies recommending policies and practices related to literacy, numeracy, and technology were excluded because each of these topics warrants its own comprehensive literature review. Research was drawn from peer and non-peer reviewed works with appropriate qualification of the author (i.e., recognized leaders in the field of education). Literature had to be published within the period of 2000-2010 (with exception for OECD sources). This period was selected because of the limited availability of empirical research on in-school and classroom practices published in journals of economics, sociology, psychology, political science and education. By searching for materials over a period of ten years, it was presumed (correctly) that a greater number of relevant articles would be located. Finally, the literature review was limited to works published in the English language however literature included in this review was published in multiple countries, including Australia, Canada, Europe (France, Germany, Luxembourg, and the United Kingdom), Hong Kong, New Zealand, and the United States.

The second step was the literature search. Four different approaches were used to locate relevant literature: 1) online search portals (Scholars Portal, Google Scholar), 2) the library of the Ontario Ministry of Education, 3) key article reference searches and 4) existing OECD literature. Over 2,000 items were located through searches and a detailed search log was maintained. For online searches using Google Scholar, the first 200 items were reviewed out of thousands of possible items. For Scholars Portal, all items located were reviewed (between 50 and 500 items per search). A combination of different search terms was used when searching for articles. Specifically, searches were conducted using the discrete terms of “practice”, “policy”, “failure”, “equity”, “education” and one of the following composite terms: intervention, leadership, management, schools, programs, assessment, and curriculum. These same terms were used to search the Ontario Ministry of Education online library index; locating a dozen additional books not identified using the online searches.

Step three involved an analysis of the materials found through the search. A review of the abstract or summary of each of the 2,000 plus items of interest was conducted based on the inclusion/exclusion criteria (described above). This step served to cull the items down to approximately 50 relevant resources. In step four, the final material was read and analysed.

ANNEX C: HATTIE'S EFFECT SIZE BAROMETER

Hattie (2009) examined over 800 meta-analyses, which encompassed 52,637 studies, based on tens of millions of students and provided 146,142 effect sizes about influence of some program, policy or innovation on academic achievement in schools (p. 15). Based on this work, Hattie provided a useful framework for interpreting effect sizes. Hattie writes:

> *"Raising achievement that is enhancing learning beyond an effect size of d = 0.0 is so low a bar as to be dangerous and is most certainly misleading. Set the bar at d = 0.40. The average effect size [across the meta-analyses] is d = 0.40 [and]…teachers typically can attain between d = 0.20 to d = 0.40 growth per year – and that is to be considered average. The effect size of 0.40 sets a level where the effect of innovation enhance achievement in such a way that we can notice-real world differences, and this should be a benchmark of such real-world change. Effects lower than d = 0.40 can be regarded as in need of more consideration, although…it is not as simple as saying that all effects below d = 0.40 are not worth having (it depends on costs, interaction effects and so on)" (Hattie, 2009, p. 16-17.)*

The book makes reference to effect sizes on hundreds of different practices mirroring the major themes identified through the literature search. These effect sizes and others are mentioned throughout the literature review as one of many sources of information in weighing the impact of various approaches on reducing school failure. The hinge point of *d* = 0.40 moves the dsicussion beyond asking "what works" to asking the question "what works best" (Hattie, 2009, p. 18).