Does Alberta Have a High-Quality Education System?

Issue:

- There is a belief, popular in some circles, that Alberta's education system in general and/or the Calgary Board of Education (CBE) in particular is broken and needs a big shake-up.¹ See example in black box at right.
- In other circles, there is a belief that while Alberta's education system – like every single education system in the world – benefits from continuous improvement, Alberta in general and the CBE in particular continues to offer one of the best education systems in the world.² See example with white background, below right.
- Which side of this debate is correct?



So far, I've come across three methods for settling this debate:

- Gut feel.
- National and international standardized tests.
- Provincial standardized tests.

I don't like gut feel. Tell me about standardized tests.

- Assessment is used to find out what students know about the subjects they study.
- Ongoing (also known as formative)
 assessment includes when teachers gather
 data via anecdotal notes, observations,
 conversations, and perhaps mini quizzes and
 tests. Some of the ways it can be collected is
 through writing samples, and conversations

2018 PISA results: Minister LaGrange

December 4, 2019



Minister of Education Adriana LaGrange issued the following statement in response to the release of the 2018 Program for International Student Assessment (PISA) results:

"Today's release of the 2018 PISA results for 15year-old students reaffirms that Alberta continues to provide our students with a high-quality education. Congratulations to participating students and teachers on these results.

"2018's results place Alberta third in the world in science, third in reading and eighth in math. This is compared to second in the world in science, third in reading and 14th in math in 2015.

One of many examples: https://twitter.com/adrianalagrange/status/1376628063186784259?lang=en

² One of many examples: https://educationnewscanada.com/article/education/level/k12/3/804548/2018-pisa-results-minister-lagrange-

 $[.]html\#: \cong : text = December \% 204\% 2C\% 202019 \& text = \% 222018 's\% 20 results \% 20 place \% 20 Alberta \% 20 third, or \% 20 see in g\% 20 a \% 20 slight \% 20 decline.$

- with students. This happens in the classroom all the time.
- Summative assessment includes standardized tests such as Programme for International Student
 Assessment (PISA), Trends in International Mathematics and Science Study and Progress in
 International Reading Literacy Study, (TIMSS and PIRLS), and Alberta's Provincial Achievement Tests
 (PATs), as well as tests administered at the end of a unit or topic of student to help the teacher
 gather evidence to support what has been noted during ongoing and formative assessment
 gathering.
- Both formative and summative assessments are important in education.
- Standardized testing can be beneficial in that it offers a benchmark to compare students' achievement. An example of this is looking at the results of Grade 6 Provincial Achievement Tests one year, and comparing it year by year at the same school to monitor for trends not only in grade 6 but in the grades leading up to it. When Grade 3 PATs were being used, K-6 schools would look to see where Grade 3 students performed well, as well as whether a student who performed well in Grade 3 also performed well in Grade 6.

So, what do international and national standardized tests say about Alberta?

- APPENDIX A has links to the originals, so you can read them in full for yourself.
- In short, although scores remain above international averages, Alberta students' results in math and science have slipped since the tests were introduced during the past 20 years. Canada's rank has slipped, too.

In other words, both sides can find proof to support their positions.

- Two people can look at standardized test results and derive different conclusions:
 - The glass-is-half-full perspective comes from focussing in on the fact that Alberta's PISA scores remain above international averages. The latest PISA results place Alberta students third in the world in science, third in reading, and eight in math. We are near the top. There is room for improvement, yes, but these rankings fall far short of the bar needed to indicate that we need to discard our current education approach and start from scratch.
 - The glass-is-half-empty perspective comes from focussing in on the PISA trends over time. Alberta has declined in science since 2015 along with Canada as a whole, and Alberta has declined in math since 2012 along with Canada as a whole. (When it comes to reading, Alberta has stayed pretty flat since 2009, but keep in mind that since Alberta ranks near the top of the world in reading, remaining flat is a good thing.) Furthermore, in the 2019 TIMMS for Grade 4, the average mathematics score for Alberta, Manitoba, and Newfoundland and Labrador is significantly below students in Canada overall. It has steadily fallen each testing period since 1995.
- So both sides are correct in this debate. To paraphrase a gem I came across in a university statistics textbook a long time ago, the competing sides can use the results of standardized tests like a drunk man uses a lamp post for support rather than illumination.

What do provincial standardized tests say about Alberta?

 Alberta collects reams and reams of data from its own students. The result is something called "accountability pillars."

Student Learning Achievement (Grades 10-12)

Diploma Exam Results By Course

Province: Alberta

English Lang Arts 30-1

	Alberta					
	2015	2016	2017	2018	2019	
Students Writing	28,166	29,730	30,150	30,393	29,832	
Diploma Examination Acceptable Standard (%)	86.5	86.8	86.5	87.5	86.8	
Diploma Examination Standard of Excellence (%)	11.4	10.7	11.7	13.2	12.3	
School Awarded Acceptable Standard (%)	97.5	97.8	97.9	97.9	98.0	
School Awarded Standard of Excellence (%)	31.9	32.2	32.9	34.5	35.1	

 They're not exciting. Students' results across Alberta have not swung wildly up and down in the last five years. (The comparison between divisions, and results of Indigenous students, and English language learners are other stories.)

There are limitations to how much we can learn from standardized assessments.

- While both formative and summative assessments are important in education, issues can arise with standardized testing when these assessments are viewed as the *only* way to understand student progress and achievement.
- Here is a list of some of the limitations to consider when evaluating the validity and usefulness of standardized tests such as PISA, TIMSS, and Provincial Achievement Tests (PATs):³
 - Two of the tests include raw scores, which people often misinterpret, because they don't realize that the raw scores are significant only to the ordinal placement of that particular exam. I take a deeper dive into this issue in APPENDIX B.
 - As many Alberta schools move to be more inclusive of students with learning disabilities, behavioural problems, cognitive disabilities and other challenges, other countries are moving away from that model, leaving the systems incomparable.
 - One of the concerns critics have about standardized tests, local or international, is how well they accommodate for the test writer possibly not having English as their first language. Are

³ Sources: https://edmontonjournal.com/news/local-news/teachers-association-wants-alberta-out-of-gamed-international-exams; https://edmontonjournal.com/news/politics/keeping-score-pisa-results-has-alberta-top-of-the-class-in-reading-science

- schools, regions, jurisdictions, with higher ELL and immigrant populations disadvantaged in rankings?
- There are concerns about the accuracy of scores from some countries with high-performing PISA results. Critics question whether all countries are following the rules in how they randomly select which pupils write the tests.
- Standardized tests alone do not effectively test whether the student understands knowledge in a real way. Consider the difference in evaluating a student on the basis of a multiple-choice exam in science compared to evaluating a student on the basis of laboratory work in science – both have their place, but hardly anyone would argue that science knowledge could be adequately tested on the basis of multiple-choice tests alone.
- Over-reliance on standardized testing results can cause a shift towards "teaching to the test", with a shift of attention to short-term fixes designed to help a jurisdiction quickly climb the rankings, rather than developing the potential of the whole child. Critics charge that PISA has led many governments into an international competition for higher test scores. Critics also charge that countries are revamping their education systems just to improve their standings with changes that don't benefit students, resulting in more and longer batteries of multiple-choice testing to practice for the tests, more scripted vendormade lessons designed to ace the tests, and less autonomy for teachers.
- By emphasising a narrow range of measurable aspects of education, standardized tests take attention away from the less measurable or immeasurable educational objectives, thereby narrowing our collective imagination regarding what education is and ought to be about.
 See APPENDIX C for a deep dive into this concern.
- The direct and indirect costs of administering PISA are substantial. The Alberta government pays around \$300,000 for a random selection of students to participate in two of the exams. In 2016, the federal government paid nearly \$7 million in fees for 22,000 students from across the country to write PISA alone. It's fair to question whether taxpayers are getting their money's worth.
- The aggregated test results, released long after students write the exams, tell teachers nothing they don't already know about their students.
- Some additional, specific limitations of PATS:
 - We know that the results of these types of assessments have less to do with the school, and more to do with the surrounding community at large.
 - Students who are ill or opt out of PATs are given a mark of 0.
- Some additional, specific limitations with PISA:
 - They do not allow for comparison at the school level because PISA is only offered to 15-year-olds, with no follow-up with the same students.
 - They use a random sample, rather than testing the entire population, and samples are rarely truly random.
 - Reading was the main focus of the 2018 test, with math and science forming only minor parts, which decreases the value of the math and science rankings for critics.
 - To carry out PISA and a host of follow-up services, OECD has embraced "publicprivate partnerships" and entered into alliances with multi-national for-profit

companies, which stand to gain financially from any deficits—real or perceived—unearthed by PISA. Some of these companies provide educational services to school districts in various PISA-participating countries on a massive, for-profit basis.

- Some additional, specific limitations with TIMSS:
 - The most recent TIMSS test only included 4% of grade 4 students in the province, and the year before that, there is a statistical note on the results from Alberta that the population was oversampled. One argument against of oversampling is that it artificially can bring down the results.
 - There's also no Grade 8 TIMSS test to compare to.

Wowzers. Let's try gut feel instead. What do our guts say about the quality of education in Alberta?

- If we use gut feel as the basis for analysis, the verdict is clear: the education system is broken. Toss it out and start from scratch!
- Canadians were asked their views on science, technology, engineering and math (STEM) education for a September 2020 Pew Research Center report, <u>Science and Scientists Held in High Esteem Across Global Publics</u>. Researchers found that while 49 percent of respondents think STEM education in Canada's universities is the best in the world or above average, for primary and secondary schools, the number fell to only 36 percent. (Whereas, if everyone knew of and trusted in the international standardized tests results, that 36% percent would have been 100%. That is a significant delta between the actual rankings compared to the perceptions of rankings.)
- An astonishing 60 percent of respondents think the standard of STEM teaching in Canada's primary and secondary schools is just "average" or "below average".
- This reflects a huge and very real lack of trust in Canada's education systems.

To clarify, are we talking about Alberta in general or the CBE specifically?

- While the two are sometimes conflated, this briefing note is focused on Alberta's education rankings in general.
- However, because the Calgary Board of Education (CBE) is the largest school board in Alberta, the
 results also reflect trends within CBE schools. One out of every six students in Alberta attends a CBE
 school.

Can we use the information covered in this briefing note to evaluate the draft curriculum proposed by the provincial government in March 2021?

No, we cannot. The results of previously conducted standardized tests contribute zero information
about the draft curriculum released in March 2021, and therefore they cannot be used to evaluate
the draft March 2021 curriculum. See APPENDIX D.

Does Alberta need an updated curriculum?

- Yes, absolutely.
- There are no "two sides" or debate on the issue of whether Alberta needs an updated curriculum.
 This is a rare point of agreement among everybody politicians across the spectrum, the Alberta Teachers' Association (ATA), school boards, teachers, parents.
- This is why the provincial Progressive Conservative Party kicked off a curriculum update in 2008.
- There are three reasons why an update is needed:
 - To incorporate the latest research about how children learn. Curriculum experts and education researchers continually learn about better ways to reach more students, and because the last major update took place a few decades ago, Alberta's curriculum was past due for an update.
 - To better reflect changes in our world. Technology, scientific knowledge and society continually evolve, and curriculums need occasional updates to remain up-to-date.
 - To reflect parental concerns. I recall when my eldest was in Kindergarten, there were parents in older grades who were concerned about aspects of the math curriculum. It came up a lot at parent council meetings back then. So the Progressive Conservative Party's 2008 kickoff of the curriculum update was also partly done to respond to concerns raised by voters.
- Prior to 2008, the curriculum was a fluid system which was continually changed and adapted, on a subject-by-subject basis. The whole-of-curriculum approach to updating adopted in 2008 was taken because subjects and skills are increasingly interconnected. For example, to increase the opportunities for drilling math basics, the latest research recommends surreptitiously incorporate math drills into other subject areas, and vice versa.
- Curriculum updates should be primarily driven by subject-matter experts, and not by politicians.

Okay, so come on now – the current education system...Broken? Or not broken?

Not broken!

...

- Alberta in general (and the CBE in particular) continues to offer one of the best education systems in the world.
- From what I've observed, Alberta's Grade 12 graduates continue to be accepted to and excel at the
 world's best universities, graduates of Alberta's schools continue to be sought-after candidates in
 the job market, and Alberta high-school graduates continue to contribute to the prosperity of our
 province as community builders, innovators, and entrepreneurs. Let's use our common sense to
 make this call!

Prepared by: 4 Susan Vukadinovic

Date: May 1, 2021

⁴ To clarify, "prepared by" is not the same as "written by". This briefing note collates observations and analysis from various people, including Janet French, Gerry Varty, and various other people who are much more knowledgeable about these issues than me. Brilliant insights are theirs; errors are mine.

APPENDIX A: Results of Four Different National and International Tests

1. PISA (Program for International Student Assessment)

PISA evaluates the performance of 15-year-old students irrespective of what grade level they are in. The majority of Alberta students who write are in Grade 10, with some in Grade 9. This assessment is carried out every 3 years since 2000, with the major study domain rotating between reading, mathematics and science (in that order) for each cycle. The latest report (PISA 2018) was released in December 2019.

In 2018 PISA focused on reading skills. Mathematics and science were assessed as minor domains, global competence as an innovation domain, and financial literacy as an optional domain.

Canadian results PISA 2018 Study: PISA2018 PublicReport EN.pdf (cmec.ca)

Insights and Interpretations 2018 PISA: https://www.oecd.org/pisa/PISA 2018 Insights and Interpretations FINAL PDF.pdf

Canadian Performance in PISA (C.D. Howe Institute) discussing trends from 2000-2018: https://www.cdhowe.org/sites/default/files/attachments/research_papers/mixed/Commentary_576.pdf

2. TIMSS (Trends in Mathematics and Science Studies)

TIMSS evaluates the performance of grade 4 and grade 8 students. This assessment is carried out every 4 years since 1995. Latest report (TIMSS 2019, Grade 4) released Dec 2020. *Minimum average age of 9.5 years old (Grade 4) or 13.5 years old (Grade 8) at the time of testing,*

Alberta did not participate in the 2003, 2015 or 2019 assessments for Grade 8.

2019 TIMSS Full Report: https://timss2019.org/reports/

2019 TIMMS Canada report:

https://www.cmec.ca/Publications/Lists/Publications/Attachments/417/TIMSS19 Report EN.pdf

3. PIRLS (Progress in International Reading Literary Studies)

PIRLS evaluates reading achievement at the 4th grade. This assessment is carried out every 5 years since 2001. Latest report PIRLS 2016. The next PIRLS assessment will be in 2021.

2016 PIRLS results: https://nces.ed.gov/surveys/pirls/pirls2016/tables/pirls2016 table01.asp

NOTE: TIMSS and PISA test different things. TIMSS (and PIRLS) is grade-based and more curriculum-based, whilst PISA is age-based and assesses the application of skills to real life problems. For a short comparison overview: https://ioelondonblog.wordpress.com/2017/12/04/how-similar-are-the-pisa-and-timss-studies/.

4. PCAP (Pan-Canadian Assessment Program)

PCAP assesses grade 8/Secondary II (Québec) students across Canada every three years since 2007. The major study domain rotates between reading, mathematics and science. In 2019 the major study domain was mathematics. *Students in French Immersion wrote the assessment in English.* Baseline years are 2007 for reading, 2010 for mathematics and 2013 for science. <u>Note</u>: 2019 PCAP results have not yet been published.

2016 PCAP report (released April 30, 2018) with a major domain of reading: https://cmec.ca/Publications/Lists/Publications/Attachments/381/PCAP-2016-Public-Report-EN.pdf

**To review results of PISA, TIMMS, PIRLS and PCAP studies during other assessment years: https://www.cmec.ca/322/International_Computer_and_Information_Literacy_Study_(ICILS)_.html

APPENDIX B: Understanding raw scores in PISA

Two of the tests (PISA and TIMMS) include raw scores, which people often misinterpret because they don't realize that the raw scores are significant only to the ordinal placement of that particular exam.

Below is a screen grab of raw scores in PISA.

Performance in reading, mathematics and science: provinces in international context (2018)

	ading		Math		Science	
S-J-Z (China)	555	B-S-J-Z (China)	591	B-S-J-Z (China)	590	
gapore	549	Singapore	569	Singapore	551	
Iberta	532	Macao (China)	558	Macao (China)	544	
cao (China)	525	Hong Kong (China)	551	Viet Nam	543	
ng Kong (China)	524	Quebec	532	Alberta	534	
intario	524	Chinese Taipei	531	Estonia	530	
onia	523	Japan	527	Japan	529	
nada	520	Korea	526	Finland	522	
land	520	Estonia	523	Quebec	522	
uebec	519	Netherlands	519	Korea	519	
ritish Columbia	519	Poland	516	Ontario	519	
and	518	Switzerland	515	Canada	518	
ova Scotia	516	Ontario	513	Hong Kong (China)	517	
ea	514	Canada	512	British Columbia	517	
fld & Lab	512	Alberta	511	Chinese Taipei	516	
and	512	Denmark	509	Poland	511	
eden	506	Slovenia	509	New Zealand	508	
w Zealand	506	Belgium	508	Nova Scotia	508	
ited States	505	Finland	507	Slovenia	507	
t Nam	505	British Columbia	504	Nfld & Lab	506	
ted Kingdom	504	Sweden	502	United Kingdom	505	
an	504	United Kingdom	502	Netherlands	503	
stralia	503	Norway	501	Germany	503	
	503		500		503	
nese Talpei rince Edward Island		Germany Ireland	500	Australia United States	503	
nmark	503 501	Czech Republic	499	Prince Edward Island	502 502	
	499	Austria	499	Saskatchewan	502 501	
rway	499	Latvia	496	Sweden	499	
askatchewan	499 498				499	
rmany		Viet Nam	496	Belgium South Bookhille		
wenia	495	France	495	Czech Republic Ireland	497 496	
Aanitoba	494		495			
Igium	493	New Zealand	494	Switzerland	495	
ince	493	Nova Scotia	494	France	493	
rtugal	492	Portugal	492	Denmark	493	
ech Republic	490	Australia	491	New Brunswick	492	
lew Brunswick	489	New Brunswick	491	Portugal	492	
CD average	487	OECD average	489	Norway	490	
therlands	485	Nfld & Lab	488	Austria	490	
stria	484	Russia	488	Manitoba	489	
itzerland	484	Italy	487	OECD average	489	
patia	479	Prince Edward Island	487	Latvia	487	
via	479	Slovak Republic	486	Spain	483	
ssia	479	Saskatchewan	485	Lithuania	482	
ly	476	Luxembourg	483	Hungary	481	
ingary	476	Manitoba	482	Russia	478	

Unfortunately, even Alberta's Minister of Education, The Honourable Adriana LaGrange, has
demonstrated a lack of awareness of how the raw scores work. The Honourable Adriana LaGrange
issued a press release on December 4, 2019 about the raw scores for PISA that said, "While these
rankings are something to applaud, the raw scores are either flat or seeing a slight decline. In 2018,
Alberta scored 534 in science, 532 in reading and 511 in math. In 2015, Alberta scored 541, 533 and
511 respectively."

- However, and this is key, the PISA tests change every time, so the raw scores are significant only to the ordinal placement of that particular exam.
- PISA has no maximum score. Scoring "511 in math", for example, is not "511 out of x". PISA is created as a standardized open-ended exam that has an arbitrary middle of 500 points by design. A score of 511 is only relevant on that particular exam in comparison to what others scored on that particular exam (i.e., did they score higher or lower than you did.)
- The exams are linked for comparison by a form of equating that indicates which exams are harder and provides a scaled approach to compare the raw scores, so someone interpreting the raw scores would have to apply the scale factors before comparing them.
- The standard deviation of a PISA exam is around 100. The standard deviation on a more familiar exam (for example, the PAT or Alberta's Grade 12 Diploma Exams) is around 20. This is handy if you are trying to figure out how significant it is that Alberta scored 511 and Quebec scored, for example, 531. If you are trying to understand what this 20-point differential means, for practical purposes, you can divide the Std Dev (100/20 = 5) to give you a rough conversion factor, which means a 20-point differential is equivalent to somewhere around 4%. So you can infer that Quebec beat Alberta by 4%. In terms of whether this differential is statistically significant depends on whether a 4% differential is maintained across time.
- This is why a math score of 511 produced a 14th-place ranking in 2015, whereas the same 511 score in 2018 produced an 8th-place ranking.
- Scores on these instruments can provide us with information that we can use to improve. However, it's important to understand what they actually *mean* in order to draws inferences from them.

APPENDIX C: Just because it's immeasurable doesn't mean it's fluffy

- By emphasising a narrow range of measurable aspects of education, standardized tests take
 attention away from the less measurable or immeasurable educational objectives like physical
 development, moral development (e.g., integrity), civic development, artistic development, and
 the development of growth mindsets, thereby dangerously narrowing our collective imagination
 regarding what education is and ought to be about.
- If you're not sure what I mean by "growth mindset", here is one example:
 - Parents today expect schools to cultivate an entrepreneurial mindset, which can't easily be
 measured through a standardized test. Cultivating an entrepreneurial mindset empowers
 kids to remain adaptable when facing obstacles, persist through failure, communicate
 better, and become better problem solvers, opportunity finders and next-generation
 innovators who can achieve their personal and career goals.
- Here is another example of a "growth mindset":
 - Conference Board of Canada (CBoC) research has shown that while Indigenous people make up four percent of adults in Canada, less than two percent of people working in STEM occupations are Indigenous. Cultivating an Indigenous learning mindset can help address this disparity. Research published by the CBoC in December 2020 found that when educators use a culturally responsive curriculum, Indigenous students are more engaged and perform better. For instance, where Western science takes a quantitative, compartmentalized approach to understanding nature, an Indigenous learning mindset in science leans towards a more qualitative, interrelated approach. An Indigenous learning mindset in science also brings distinct views about interconnectedness, relationship to space and time, structural authority, and ways of knowing and learning. The report found that Indigenous students benefit from a more culturally appropriate approach to STEM learning.
- Other key growth mindsets we expect our schools to cultivate are: **equity**; **diversity**; and **inclusivity**.
- Our children's future success in STEM careers won't just involve technical prowess with calculations
 and coding. The most successful people in STEM careers currently and in the future are those who
 are creative, have growth mindsets, and are capable of combining evolving technology and the
 ability to collaborate with other professionals in other fields.
- Success in STEM careers also doesn't just involve instilling technical prowess with current
 calculations and coding practices. There's a massive digital transformation happening in our world
 and our education system needs to prepare our kids to succeed in an Industry 4.0 that doesn't even
 exist yet: smart digital technology, machine learning, big data, and other evolving systems that are
 changing what we know about design, logistics, optimization and life cycle.
- Our education system needs to prepare our kids not for today, but for 10 to 20 years from now. We are educating students to meet the skills gap that exists today and into the future.
- A level of comfort with the knowledge base for technology, math, science and engineering is absolutely important. The capability and skills set that our children require is exceptional, and as a society we cannot drop the ball on covering the basics.
- However, we'll shortchange our kids if we don't also nurture less measurable attributes.

APPENDIX D: Draft curriculum = raw chicken

Using the results of standardized test scores from previous years is completely irrelevant to evaluating the draft March 2021 curriculum.

Asking the internet is also not necessarily the best way to evaluate the draft March 2021 curriculum (although the meme below that I found circulating on social media is pretty funny).

Instead, if you want to truly understand the value of the draft March 2021 curriculum, please ask curriculum experts and teachers with real-world classroom experience.

Draft Curriculum = Raw Chicken (an analogy)

I once ordered a chicken burger at a diner and when it came the breast wasn't just undercooked, wasn't just raw ... it had ice crystals in the middle. Once I saw that, I sent it back.

I didn't stop to check if the bun was good, if the tomatoes and lettuce were fresh, nor if the fries were hot. I didn't even check to see if PARTS of the chicken were cooked. Neither did I conduct an analysis of the bacterial count, nor write out step by step instructions for the kitchen on how to cook chicken. I just saw a fatal flaw in what I had been served, deemed it inedible, and sent it back for them to do it again.

Think of this draft curriculum like that, except that this government has served raw chicken to an ENTIRE ROOM full of cooks (teachers), culinary instructors (education profs), and culinary scientists (subject area academics), along with many members of the general public. Almost everyone in the room has sent their meal back for a do-over.

Unfortunately, the response from the restaurant has been combative: "Just because Jason Kenney owns this diner and Adriana LaGrange runs the kitchen, y'all are trying to make a big deal out of nothing (at this point management points out a loyal patron, Mr. Staples, who is eating raw chicken with gusto). You're just politicizing this meal."

No we aren't. <u>This curricular chicken is raw, folks.</u> It doesn't matter who cooked it or who owns the joint. It also doesn't matter if other parts of the meal were A-OK. We're sending this undercooked curriculum back. Do it over. And take that burger from Mr. Staples before he winds up in the hospital, for pity's sake.