**AI, Classroom Assessment, and Authentic Learning:
Key Considerations for School System Leaders**

Louis Volante (PhD)

Distinguished Professor, Brock University

Professorial Fellow, UNU-MERIT / Maastricht University

CASSA Leadership Conference

Montreal, Quebec

July 4, 2024

AGENDA

- PART I: AI and Classroom Assessment

- PART II: Academic Resilience

- PART III: COVID, Learning Loss, and 'Catch-Up' Policies

- Group Q & A

PART I

AI and Classroom Assessment

ChatGPT

- With 100 million active users in January 2023, just two months after its launch, ChatGPT is considered the fastest-growing consumer internet application.

- This popular AI chatbot, developed by OpenAI, has precipitate a 'tsunami effect' of changes to education (García-Peñalvo, 2023).

- Users as young as 13, *with parental or guardian permission*, are now able to generate essays within seconds via the AI application, complete with references.

- Other companies have developed similar AI language models, such as Google's BARD or Microsoft's Sydney, to rival the overnight success of ChatGPT.

AI, Software Detection, and False Positives

- More than 50% of high school students in the US reported some form of cheating that could include copying an internet document to submit as part of an assignment and/or cheating during a test (Eaton & Hughes, 2022).

- Cheating in Canada is also reported by >50% of high school students, with higher percentages (73%) reported for written assignments (Eaton & Hughes).

- In both Canada and the U.S., the incidence rates for undergraduate students are significantly lower (approximately 5%), but are still a noteworthy issue.

- Detection software companies, such as Turnitin, are promising more effective results (e.g., a 97% success rate for ChatGPT authored content, TurnItIn, 2023). Nevertheless, simple adjustments can help students evade detection.

AI in the Classroom

- Estimates suggest more than 70% of teachers have not received any guidance in the use of ChatGPT (Jimenez, 2023), despite the fact that more than 50% of K-12 teachers and 1/3 of students 12-17, report using this application (Kingston, 2023).

- ChatGPT, and other AI applications, generate texts which often look sophisticated, but are prone to factual errors. This ability to distort scientific facts and spread misinformation (van Dis, Bollen, Zuidema, van Rooij, & Bockting, 2023), also presents an opportunity to develop research literacy skills.

- There is an urgent need to offer pragmatic solutions to this pressing challenge, so that teachers and students use AI in ways that support learning.
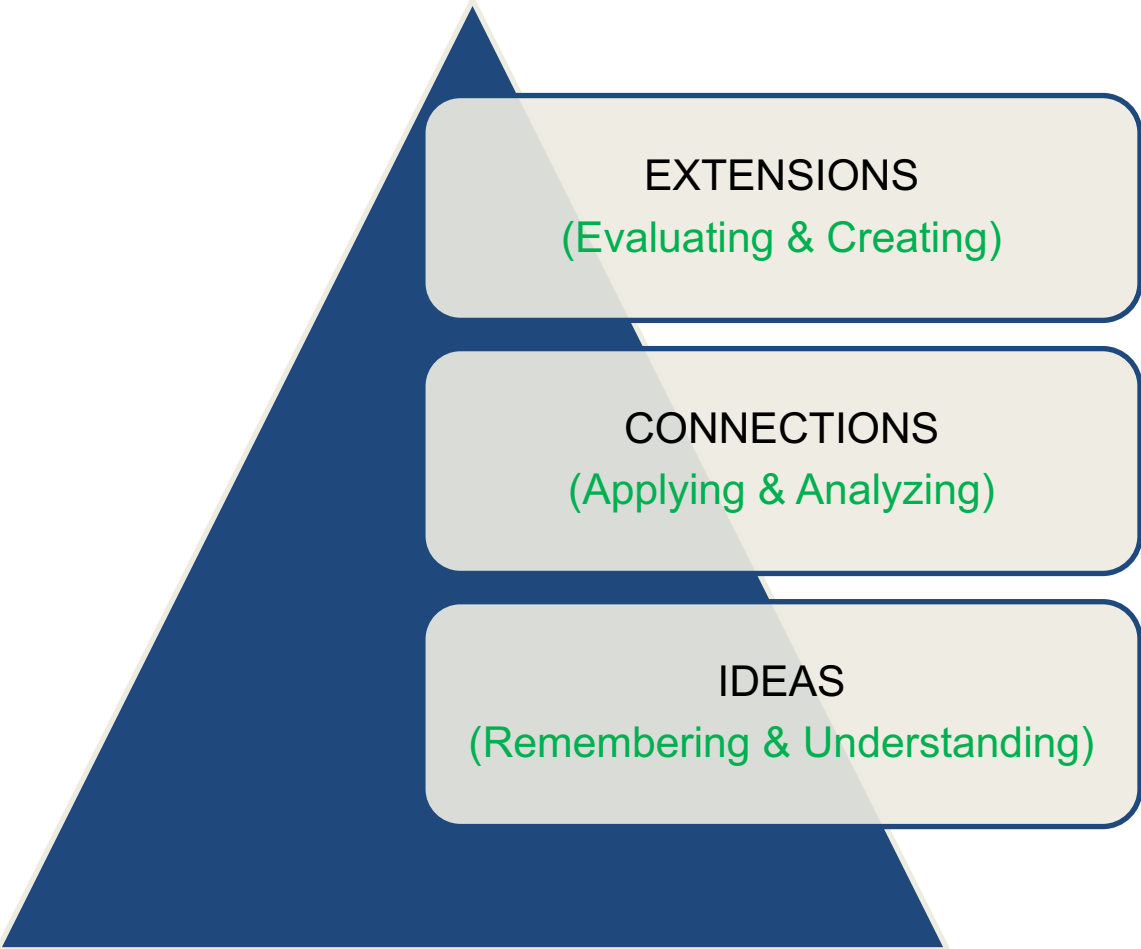
(in)Equality of Opportunity

- ChatGPT is able to pass national high school examinations (de Winter, 2023), as well as college and university law and science courses (Choi et al., 2023; Gibson et al., 2023). Simulations Results: [GPT-4 (openai.com)](#)

- School bans of AI and large language models exacerbate inequality in student learning and achievement.

- Educationally defensible uses of AI need to be considered alongside existing inequality in access to hardware, software, and connectivity issues.

- School administrators and teachers also need appropriate PD to address this elephant in our schools.

AI Generated Content

- Two critical measures of AI language models are **perplexity** and **burstiness**.

- Quite simply, perplexity relates to the *complexity of the words* in a sentence, while burstiness relates to the *variance between sentences*.

- AI language models tend to generate sentences with lower perplexity, meaning they are less complex. Conversely, humans tend to write in bursts with short and long sentences.

- Taken together, AI generated content tends to be more uniform, and less interesting, than human generated content. It is for this reason that AI content is often described as '*shallow*'.

Formative Assessment Pathway

Formative Assessment Pathway (Step 1)

- The FIRST step in using AI-generated text is for the teacher to engage students in '*fact-checking*' by referencing the key ideas in a sample output.

- After completing an initial set of revisions, students verify facts and ideas, while teachers could facilitate a peer-assessment feedback process using cooperative learning strategies.

- Overall, engaging in initial content verification step helps develop students' capabilities in relation to understanding facts, figures, and knowledge, and generally aligns with the **IDEAS** level of the ICE model.

Formative Assessment Pathway (Step 2)

- The SECOND helps students add texture to the sample writing by *revising sentence and argument structures* as well as synthesizing ideas in a more creative fashion.

- Teacher can incorporate self- and peer-review elements where students generate their own perplexity and burstiness scores (out of 10), as an engaging way to revise their work.

- This step also invites students to make connections to their personal life and local context, an action that no AI application can do.

- Collectively, this step aligns with the **CONNECTIONS** level within the ICE model, which is principally focused on linking ideas.
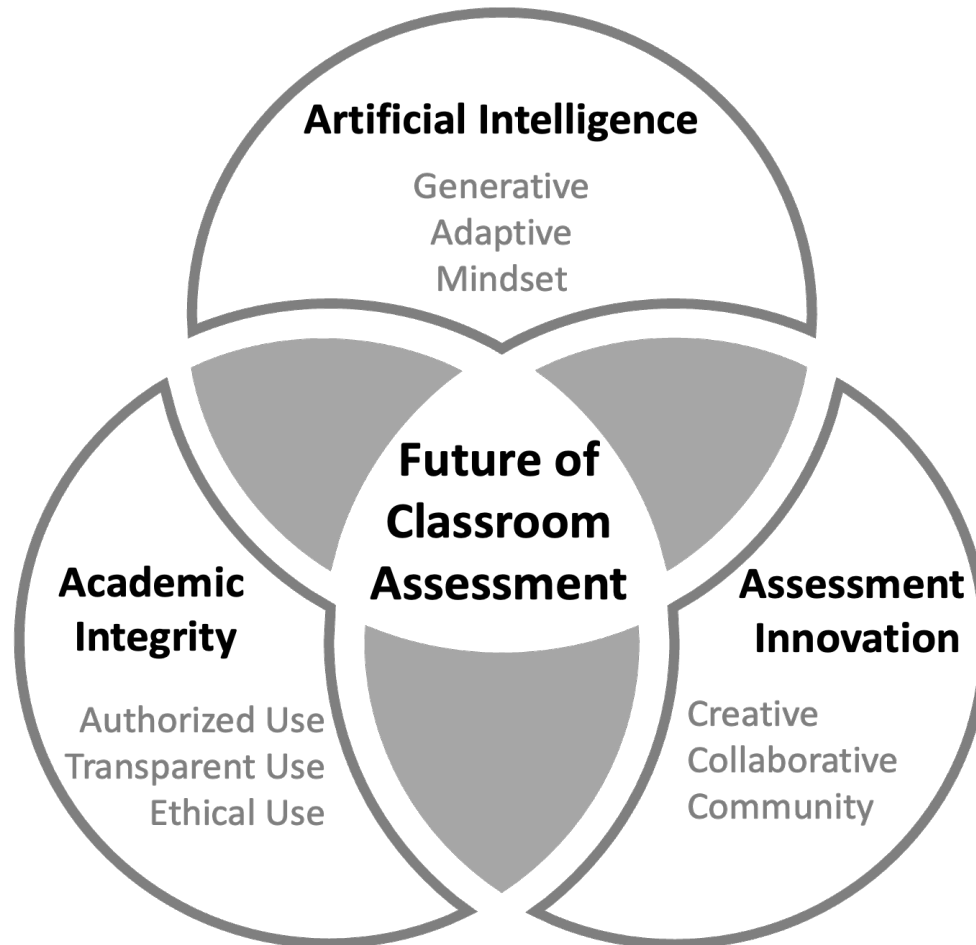
Formative Assessment Pathway (Step 3)

- The THIRD step requires students to undertake a final round of revisions to their text that illustrates **EXTENSIONS** *indicative of critical, creative, and higher-order thinking*.

- Extensions should be reflected in the assessment criteria, such as rubrics, and students should be invited to explicitly make these extensions in their work.

- Inviting students to localize extensions to their personal context (i.e., community). Provide alternative formats or authentic tasks, such as an oral presentation of the assignment, an artistic representation, or a community-based project.

- AI and the necessity of a greater emphasis on authentic assessment, provides the impetus to make education *more* human, not less (Cope et al., 2020).

Key Considerations

- Emphasize distinction between low- and higher-order thinking skills in summative assessment tasks.

- Provide observable 'look-fors' in student assessment criteria.

- Offer in-service training to highlight *selective* and *age-appropriate* incorporation of AI in classroom assignments and in authentic performance tasks.

- Clearly communicate to parents both the *opportunities* and *challenges* presented by the proliferation of AI tools on home computer systems.

AI³ Model and Forward-Looking Assessment Systems

Artificial Intelligence

Generative
Adaptive
Mindset

Future of Classroom Assessment

Academic Integrity

Authorized Use
Transparent Use
Ethical Use

Assessment Innovation

Creative
Collaborative
Community

Future-Focused Education

Can the challenge for educators change from "how do I ban or detect design tasks that AI can't do", towards how to represent a world where these tools are normal? What might we learn from previous technology panics and a history of transitioning from worry about new technologies such as writing, calculators and the Internet, to embracing them and even incorporating them into learning outcomes?

Dawson, P. (2023). *Don't fear the robot: Future-authentic assessment and generative AI.*
https://werklund.ucalgary.ca/dont-fear-robot

GenAI: Generating Administrators' Involvement - CAP (cdnprincipals.com)

"We need to be future-focused … and figuring out what our school systems need to do to ensure we are preparing students for their future and not our past."

Preparing Students for Their Future, Not Our Past | AASA

(AASA –American Association of School Administrators)

SWOT analysis is a tool for auditing an organization and its environment. It is the *first* stage of planning and can help schools/departments/grade level teams focus on key teaching and learning issues.

SWOT stands for *Strengths*, *Weaknesses*, *Opportunities*, and *Threats*. Strengths and weaknesses are internal factors. Opportunities and threats are external factors.

| Strengths	Weaknesses
Opportunities	**Threats**

Select Publications

- Volante, L., & DeLuca, C. (2024). Large-scale testing in the face of AI. *Assessment & Development Matters*.

- Volante, L., DeLuca, C., & Klinger, D. A. (2023). Leveraging AI to enhance learning and formative assessment in secondary schools. *Phi Delta Kappan*.

- Volante, L., DeLuca, C., & Klinger, D. A. (2023). Forward-thinking assessment in the era of Artificial Intelligence: strategies to facilitate deep learning. *Education Canada*.

- Volante, L., DeLuca, C., & Klinger, D. A. (2023). How can teachers integrate AI within schools? Five steps to follow. *Education Canada – The Facts on Education Series*.

PART II

Academic Reslience

Multilevel Framework
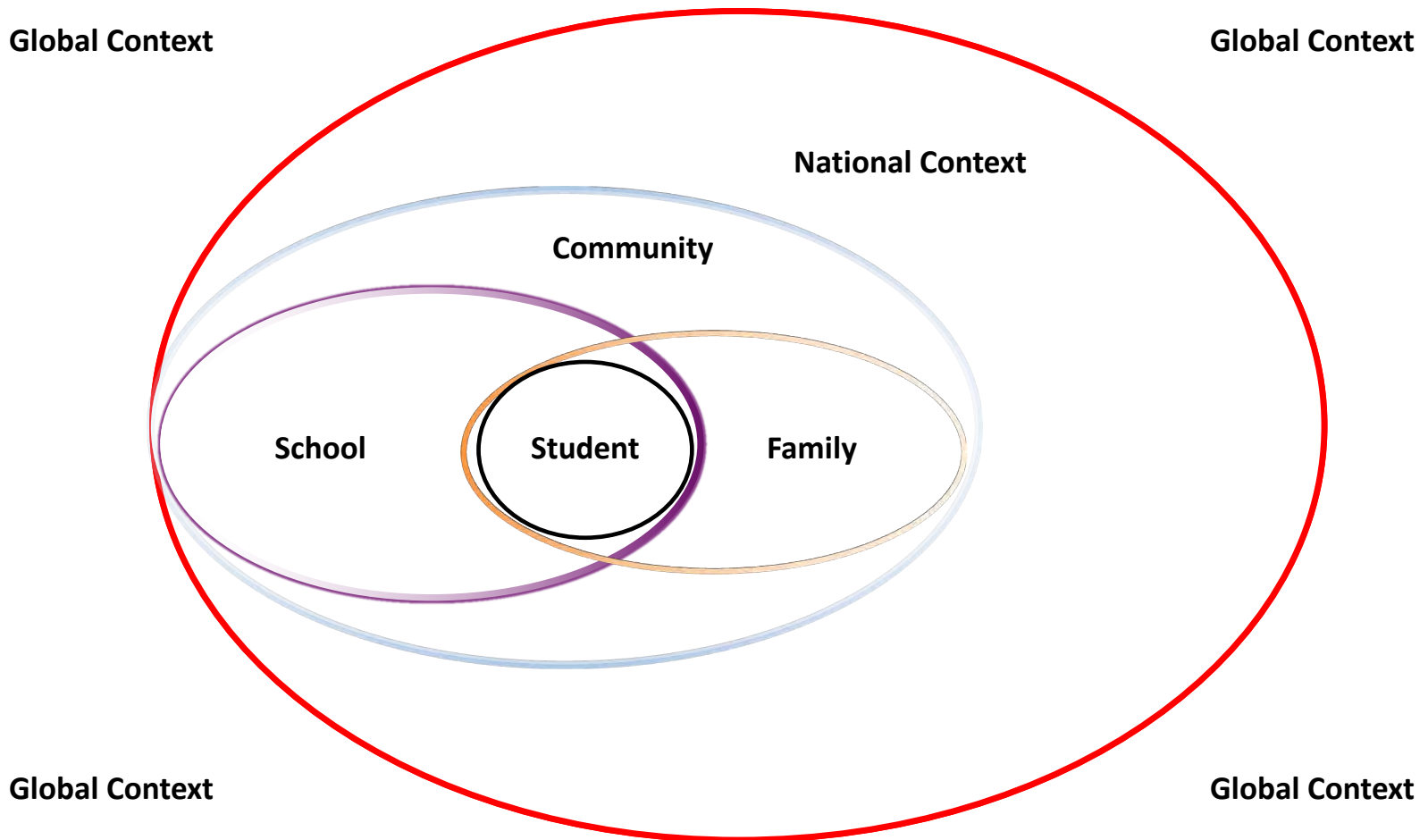
'Academic Resilience'

- Academic resilience is the general notion that there are some students who achieve favourable achievement outcomes despite coming from lower SES backgrounds.

- Countries which possess a higher relative share of low SES students who achieve well are said to have a more *academically resilient* population.

- The Organization for Economic Cooperation and Development (OECD) Programme in International Student Assessment (PISA) is increasingly the key measure used to make cross-national comparisons of academic resilience.

Cognitive, Non-Cognitive, & Interdisciplinary Skills

- The COVID-19 pandemic has forced us to reconsider the relative importance ascribed to *cognitive* versus *non-cognitive* skills, an admittedly problematic albeit common distinction in the assessment literature.

- Cognitive skills are typically measured in relation to long-standing content areas such as reading, mathematics, and science literacy which are captured by provincial, national, and international LSA's.

- Non-cognitive and interdisciplinary skills are also being increasingly measured cross-nationally (i.e., PISA) and include measures such as digital literacy, creativity, growth mindset, and socioemotional learning.

Mental Health and Physical Wellbeing

- Over 90% of enrolled learners (1.5 billion young people worldwide) had their education disrupted due to the pandemic.

- School closures contributed to increased incidence rates of depression, anxiety, irritability, attention, hyperactivity, and obsessions/compulsions in school-aged children.

- Similarly, only a fraction of children and adolescents are meeting recommended physical health guidelines during and after the pandemic.

- These trends are also found in young adults attending post-secondary education institutions.

Correlates of Student Achievement

Various *individual, family, school, community, national,* and *global* characteristics are associated with (under)achievement such as …

- Gender
- SES
- Parental education levels
- Access to early childhood education
- Age and country of arrival for immigrants
- Regional and school demographics
- *Teaching and learning environments*
- *School system features*
- Community support programs
- Income inequality of society
- Social protection and welfare provisions
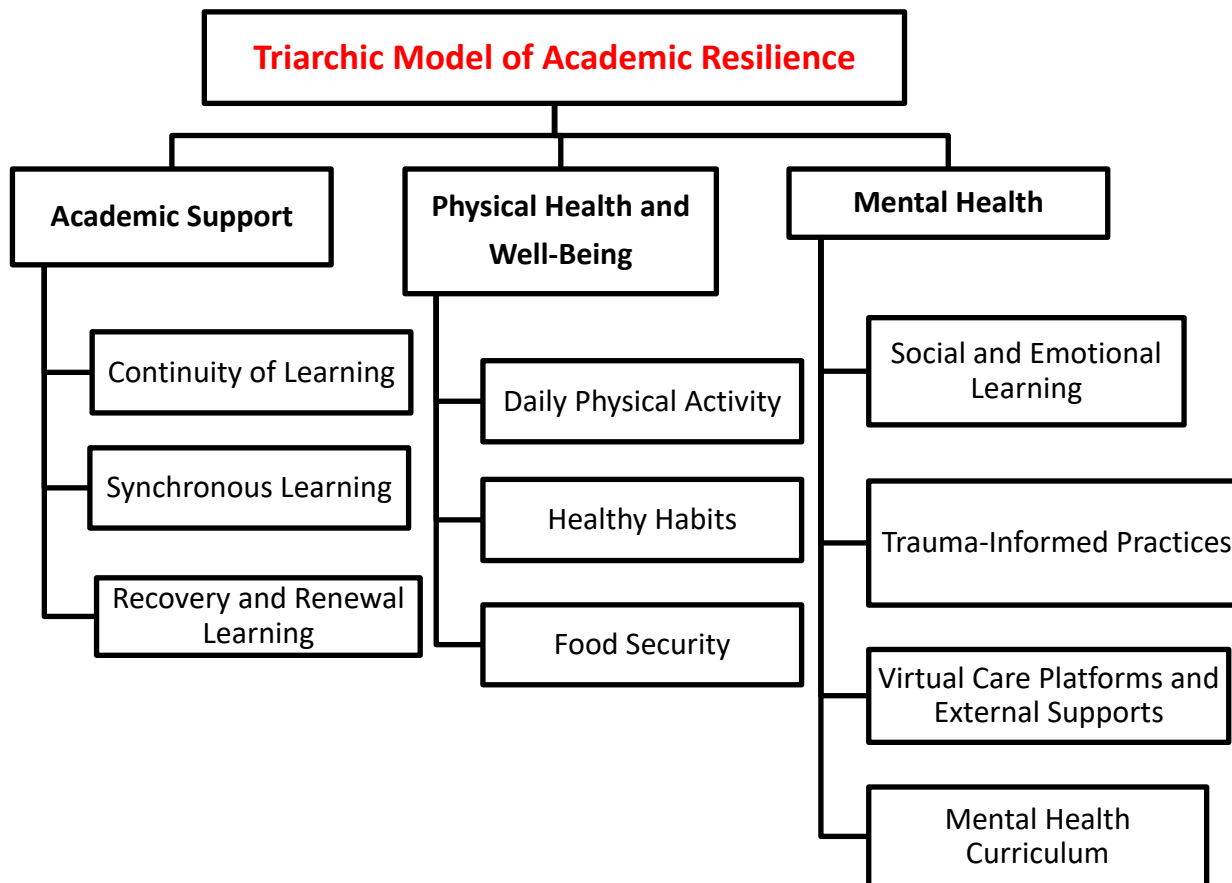- Pandemic disruptions and related policies

Education Policies

Each of the following is associated with *lower* educational outcomes:

- Lack of provisions for ECE

- Grade repetition

- Early tracking

- School choice

- Non-weighted funding formulas

- Rigid primary-secondary pathways

- Reductionist 'catch-up' COVID-19 policies

Canadian Case Study

Triarchic Model of Academic Resilience

Academic Support
- Continuity of Learning
- Synchronous Learning
- Recovery and Renewal Learning

Physical Health and Well-Being
- Daily Physical Activity
- Healthy Habits
- Food Security

Mental Health
- Social and Emotional Learning
- Trauma-Informed Practices
- Virtual Care Platforms and External Supports
- Mental Health Curriculum

Select Publications

- Schnepf, S., Volante, L., Klinger, D. A., Giancola, O., & Salmieri, L. (Eds.) (2024). *The pandemic, socioeconomic disadvantage, and learning outcomes: cross-national impact analyses of education policy reforms*. Publications Office of the European Union.

- Volante, L., & Klinger, D. A. (2023). PISA, global reference societies, and policy borrowing: The promises and pitfalls of academic resilience. *Policy Futures in Education*.

- Volante, L., & Klinger, D. A. (2023). COVID-19 and the learning loss dilemma: the danger of catching up only to fall behind. *Education Canada.*

- Volante, L., Lara, C., Klinger, D. A., & Siegel, M. (2022). Academic resilience during the COVID-19 pandemic: a triarchic analysis of education policy developments across Canada. *Canadian Journal of Education*.

PART III

COVID Learning Losses

PISA 2022: English-Speaking Countries

	Mathematics	Reading	Science
Canada	497 (1) (Global = 6th)	507 (2) (Global = 6th)	515 (1) (Global = 5th)
Ireland	492 (2)	516 (1)	504 (T-3)
Australia	487 (4)	498 (5)	507 (2)
New Zealand	479 (5)	501 (4)	504 (T-3)
UK	489 (3)	494 (6)	500 (5)
USA	465 (6)	504 (3)	499 (6)
OECD	472	476	485

Canadian Performance: PISA 2022 vs. PISA 2018

	Mathematics	Reading	Science
PISA 2022	Canada = 497 (OECD = 472)	Canada = 507 (OECD = 476)	Canada = 515 (OECD = 485)
PISA 2018	Canada = 512 (OECD = 489)	Canada = 520 (OECD = 487)	Canada = 518 (OECD = 489)
CDN Performance Decline	Drop of 15 points	Drop of 13 points	Drop of 3 points
OECD Performance Decline	Drop of 17 points	Drop of 11 points	Drop of 4 points

Learning Losses Captured by PISA

- $Y_i = \alpha + \beta_1 COVID + \beta_2 Time\ trend + \delta_i X_i + \alpha_i + \varepsilon_i$

- Y_i represents the outcome variables: For example, the standardized index of the sense of belonging to school.

- X_i is a set of student-level controls, specifically: the socioeconomic status index, gender, migrant status.

- Two country-level controls are also included:
(i) the number of weeks of full closure reported by UNESCO; and
(ii) the growth rate of spending on secondary education provided by OECD.

# Key Findings: Non-Cognitive Domain

- In relation to past cohorts, students in the 2022 PISA experienced:
  - Loss of 0.07 SD in the sense of belonging at school;
  - Increase of 0.14 SD in fixed mindset;
  - 3% decrease in probability of being frequently bullied.

- Girls, migrants, and economically disadvantaged students notably suffered greater losses in their sense of belonging than their peers.

- The pandemic exacerbated instances of bullying among girls and prompted a shift towards a fixed mindset, particularly affecting female students.

Key Findings: Cognitive Domain

- COVID-related losses for students with high levels of a sense of belonging and a growth mindset were approximately:
    - One school year for reading; and
    - Half a school year for mathematics and sciences.

- For students who are frequently bullied, the COVID-related losses averaged *1.5 school years* across all three test domains: mathematics, reading, and science literacy.

- Globally, learning losses attributable to the pandemic are unprecedented in relation to cross-national achievement data spanning more than 30 years.

Time Series Trends: Sense of Belonging
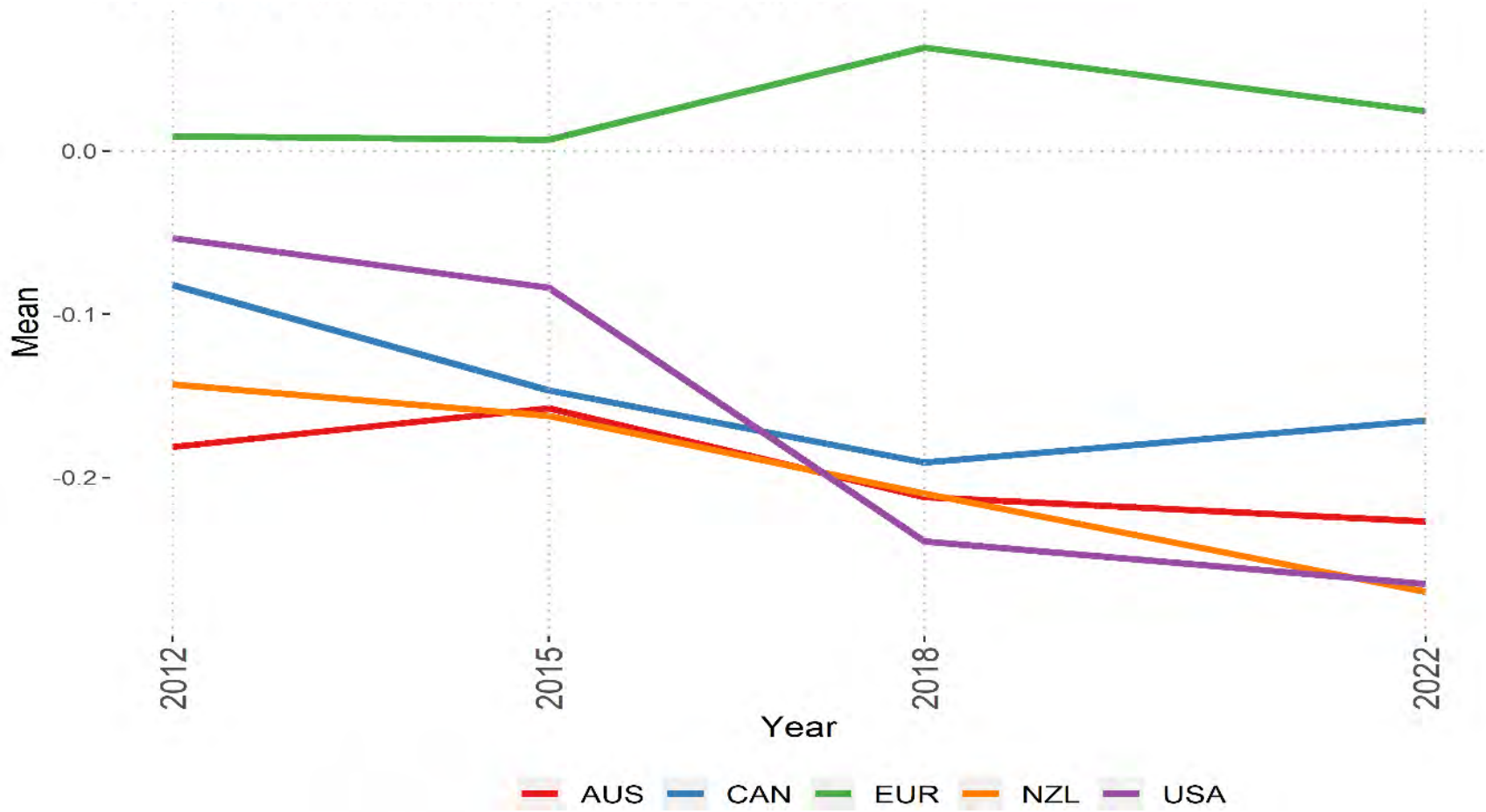
Academic Resilience & 'Catch-Up' Policies

Discuss the policies your school district and/or province is currently utilizing to support academic resilience?

Jot down a brief set of points to identify:

- Strengths

- Weaknesses

- Possible Next Steps

Finally, briefly discuss your initial thoughts on how your provincial approach is / is not aligned with the available best-practice research evidence.