SESSION SUMMARY

Standard Over Task: Re-Engineering Your Math Gradebook

Aaron Szott · Szott Education Consulting · Make Math Moments Virtual Summit 2026

Most math gradebooks are organized by product: assignments, quizzes, projects, unit tests. This summary looks at what happens when the columns in the gradebook are the learning instead: the standards, outcomes or expectations students are working toward.

It isn't a magic software trick or a fast fix. It's a different way of looking at how we arrive at a student's grade, and what that grade really communicates. My goal is simply to put a pebble in your shoe: a small idea you can't quite stop noticing. This summary accompanies my session at the Make Math Moments Virtual Summit 2026, where I walk through each example step by step.

1 Start with what you're grading

The real starting point for any school's journey toward standards-based assessment is a shared understanding of what grades are for in the first place. That's a conversation for another day. This summary picks up at the gradebook, and that part of the journey starts here.

Before you touch a gradebook, you need to be able to say clearly what students should know and be able to do by the end of the course. Whatever your curriculum calls them (outcomes, standards, expectations), that list is your starting point.

- **Know the layers.** Ontario's Grade 9 mathematics course has 5 reportable strands, lettered B to F (Number, Algebra, Data, Geometry and Measurement, Financial Literacy), 11 overall expectations and 43 specific expectations (Ontario Ministry of Education, 2021). Think of strands as your units, overall expectations as the big ideas, and specific expectations as the finer detail.
- **Pick your grain size.** How far down you track is your call. Going ultra-granular is tedious and hard to keep up with. Track what helps you make a decision, not everything you could possibly track.
- **Aim for a manageable number.** Maybe 15 to 25 standards: not so few that you lose important detail, and not so many that every question becomes a column. Some curriculum documents give you a manageable list; others (many Common Core courses, for example) will need consolidating.

Consolidate with your teaching partners and your department. AI can be a very helpful thought partner here, if you give it clear rules. The ones I used:

- **Keep neighbours together.** Combine specific expectations that sit side by side and overlap. Don't merge topics you'll teach two months apart.
- **Balance the demand.** Leave big, cognitively demanding expectations on their own; combine smaller ones so each new standard asks about the same of students.
- **Your team decides.** AI drafts. You and your colleagues review it and make the final call. (In my first pass, the AI slipped in a Grade 10 idea that isn't in the Grade 9 course. That's exactly why people check.)

In the Grade 9 example, 43 specific expectations became 24 gradebook standards: B1 to B6, C1 to C8, and so on.

© 2026 Szott Education Consulting. You may share this summary freely in its complete, unaltered form. Not to be adapted, sold or reproduced in part without written permission.

2 What a gradebook built on products hides

Imagine you're hosting a big party and you have to pick everyone up. You send every vehicle you've got: a truck, a small car, a few minivans, maybe even a horse. Nobody cares how anyone arrived; what makes the party great is the people who are there. Assessment works the same way. Tests, quizzes and assignments are the vehicles. The evidence of learning they carry is the people, and that's what matters.

A small shift is to stop asking "what type of assessment is this?" and start asking "what outcome am I gathering evidence for?"

Here is a sample student, Mateo, in a traditional Grade 9 math gradebook: five units, each weighted, and the same five products inside every unit, each weighted again. His Data unit comes out at 78%.

FIGURE 1

The gradebook we know: units, then products

Unit	Unit weight	Assignment 1 10%	Quiz 1 20%	Quiz 2 20%	Project 10%	Unit Test 40%	Mateo's unit grade
Unit 1: Number	25%	…	…	…	…	…	**74%**
Unit 2: Algebra	33⅓%	…	…	…	…	…	**76%**
Unit 3: Data	16⅔%	57	64	73	85	90	**78%**
Unit 4: Geometry and Measurement	16⅔%	…	…	…	…	…	**80%**
Unit 5: Financial Literacy	8⅓%	…	…	…	…	…	**85%**
Course grade: a weighted mean of five unit grades, each a weighted mean of five products							**77%**

Every unit holds the same kinds of products, each with its own weight. Mateo's Data unit marks are shown. Each column names the vehicle, not the learning.

What does 78% mean? That he knows 78% of the unit? That he mastered most of it and struggled with a few pieces? That he was weak at first and strong by the end? Any of these could produce a 78. The number can't tell you which.

- **A mean average only makes sense over measures of similar content.** Averaging marks from products that each assess a different mix of things is averaging apples and oranges (Canady, 1993).
- **Mean averaging everything miscommunicates.** It isn't an accurate report of what a student knows and can do at the time of reporting. Guskey (2002) asks whether a karate student who starts with a white belt and earns a black belt "actually deserves a gray belt" (p. 778).

What changes when the columns change? Here's what one school found:

> *"Some of our math teachers changed their categories to match the objectives for that quarter . . . . Kids knew that their grades were strong in linear tables and graphs but not in algorithms. So they knew that's where they needed to work. Their folks knew it, too. Before they just knew that they had low quiz grades and they were missing two homework assignments."*
>
> (Wormeli, 2006, p. 162)

© 2026 Szott Education Consulting. You may share this summary freely in its complete, unaltered form. Not to be adapted, sold or reproduced in part without written permission.

3 What a gradebook built on standards could look like

Same course, same student, same assessments. Only the architecture changes.

- **One column per standard.** Not Quiz 1, Quiz 2, Unit 1, Unit 2: the columns are the learning (Figure 4). How you weight the standards is your team's call.
- **Map every question to a standard.** You don't need new quizzes or tests. Look at the questions you already ask and record each one under the standard it gives you evidence for, whether it came from a quiz, a test, an assignment, a conversation or an observation (Figure 2).
- **Far more evidence from the same work.** Mateo's five Data assessments gave 5 numbers in the old gradebook and 28 pieces of evidence in the new one, with no extra marking.

FIGURE 2

Same five assessments, every question mapped to a standard

Standard	Assignment 1					Quiz 1					Quiz 2					Modelling Project					Unit Test							
	57% (10%)					64% (20%)					73% (20%)					85% (10%)					90% (40%)							
D1 Big data	50	60																			90							
D2 One-variable			55	60	60	65	70															95	90					
D3 Two-variable								55	60	70	70	75												90	85			
D4 Modelling													70	75	75	80	85	85	85	90						90	90	90

Each column is one question, recorded under the standard it gave evidence for. The same marking that produced five product marks now gives 28 pieces of evidence.

Once the product labels are gone, evidence is just evidence. Line it up by standard, in the order it happened (Figure 3), and you can see Mateo's journey on each one. Take D2, one-variable data: his first five marks run from 55 to 70, and his last two, on the unit test, are 95 and 90. The mean says 71 and the median says 65, but his most recent, consistent work says 90.

- **Determine, don't just calculate.** Grades should reflect the student's most consistent level of achievement, with emphasis on more recent evidence (O'Connor, 2017). The numbers can inform the decision, but they shouldn't make it (Schimmer, 2016).

FIGURE 3

Evidence is evidence: mean, median or judgment?

Mateo's Data unit	Evidence, in the order it happened	Mean	Median	Recent and consistent
D1 Big data: implications	50 60 90	66.7	60	**90**
D2 One-variable data	55 60 60 65 70 95 90	70.7	65	**90**
D3 Two-variable data	55 60 70 70 75 90 85	72.1	70	**88**
D4 Mathematical modelling	70 75 75 80 85 85 85 90 90 90 90	83.2	85	**90**

Graded by product, this unit is 78%. Neither the mean nor the median shows where Mateo is now; the evidence does.

© 2026 Szott Education Consulting. You may share this summary freely in its complete, unaltered form. Not to be adapted, sold or reproduced in part without written permission.

Those four judgments go into the course gradebook:

FIGURE 4

Into Mateo's course gradebook: one column per standard

Student	B Number						C Algebra								D Data				E Geom. & Meas.				F Fin. Lit.		Course grade
	B1	B2	B3	B4	B5	B6	C1	C2	C3	C4	C5	C6	C7	C8	D1	D2	D3	D4	E1	E2	E3	E4	F1	F2	
Mateo	–	–	–	–	–	–	–	–	–	–	–	–	–	–	90	90	88	90	–	–	–	–	–	–	–

24 columns, one for each consolidated standard, weighted equally here (1/24 each). Every column names the learning. The other columns fill in as evidence comes in, and every one stays "in pencil": open to change when newer evidence shows more learning.

The goal isn't a different grade. It's a grade that is more transparent and more defensible.

Honouring recency and consistency has nothing to do with inflating grades; no part of grading by standard does. For a student who started strong and slid, recency works exactly the same way.

4 What it makes possible

- **A question is a question is a question.** A question gives the same evidence whether it sits on a quiz or a test. Once you stop grading by product, you can quiz as you go and cycle back to earlier standards. That opens the door to far fewer heavily weighted, one-and-done test days.
- **Tiered entry points.** Carmen Sinatra teaches a markless Grade 9 math class at Richmond Hill High School in Ontario. On some of her quizzes, a few questions come in tiers: students choose two of six questions labelled Developing, Refining and Mastery. Every student gets a foothold, and every student can show how far they can go (Gogerty, 2016). Not every question needs to be tiered.
- **Grades written in pencil.** When newer evidence shows more learning, it replaces the older evidence (Guskey, 1996; Schimmer, 2016). Students can keep updating their standing on earlier topics (Marzano & Heflebower, 2011).
- **Missed assessments stop being a crisis.** First question: does it even matter? If you already have enough evidence on those standards, move on. If you need more, target just those standards. A conversation or an observation counts. You reassess the standard, not the test.
- **Students can be strategic.** When questions say which standards they assess, students who track their own evidence know where to put their effort. Carmen reports that requests for extra time have all but vanished.
- **A word about percentages.** If your school uses percentages and you want to keep using them, go ahead; many schools that grade by standard do. Grading by standard does tend to be easier with a small scale. Scales with 3 to 5 clearly described levels are more reliable and give students and parents more useful information (Link & Guskey, 2022), and some schools go up to 7. Where percentages still need to be reported, one option is to grade with a scale and convert to a percentage at the time of reporting.

© 2026 Szott Education Consulting. You may share this summary freely in its complete, unaltered form. Not to be adapted, sold or reproduced in part without written permission.

5 A pebble in your shoe

I hope this summary has you thinking about how a gradebook could be re-engineered to give us better data: data that lets us determine grades in a more defensible, fair, and transparent way for our students.

Re-engineering a gradebook is one piece of a bigger journey toward standards-based assessment and reporting. The gradebook is the part you can see. The harder parts sit underneath it: agreeing as a staff on what grades are for, consolidating a whole department's courses into standards, and building professional judgment that holds up with students, parents and administrators. Those are the parts that are hard to do alone.

How could your gradebook give you better evidence for more defensible, fair, and transparent grades?

Let's keep the conversation going

I work with individuals, departments, schools, divisions and districts on standards-based assessment and grading.

After more than 20 years as a teacher, including many years as a department head, I offer support grounded in research and experience, and tempered by empathy, for reform that is sustainable, practical and built with teachers, not handed to them.

aaron@szottassessment.com

LinkedIn www.linkedin.com/in/aaronszott

Instagram @szottonassessment

References

Canady, R. (1993, April). [Workshop presentation]. Association for Supervision and Curriculum Development Annual Conference, Washington, DC, United States.

Gogerty, J. I. (2016). *The influence of district support during implementation of high school standards based grading practices* [Doctoral dissertation, Drake University].

Guskey, T. R. (Ed.). (1996). *Communicating student learning*. Association for Supervision and Curriculum Development.

Guskey, T. R. (2002). Computerized gradebooks and the myth of objectivity. *Phi Delta Kappan, 83*(10), 775–780. https://doi.org/10.1177/003172170208301013

Link, L. J., & Guskey, T. R. (2022). Is standards-based grading effective? *Theory Into Practice, 61*(4), 406–417. https://doi.org/10.1080/00405841.2022.2107338

Marzano, R. J., & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership, 69*(3), 34–39.

O'Connor, K. (2017). A case for standards-based grading and reporting. *School Administrator, 74*(1), 24–28.

Ontario Ministry of Education. (2021). *The Ontario curriculum: Mathematics, Grade 9*. Queen's Printer for Ontario.

Schimmer, T. (2016). *Grading from the inside out: Bringing accuracy to student assessment through a standards-based mindset*. Solution Tree Press.

Wormeli, R. (2006). *Fair isn't always equal: Assessing & grading in the differentiated classroom*. Stenhouse; National Middle School Association.

With thanks to Carmen Sinatra of Richmond Hill High School, Ontario, for our valuable and insightful conversations.

This summary was prepared by Aaron Szott with drafting and layout assistance from Claude, an AI model made by Anthropic. All content was reviewed and approved by the author.

© 2026 Szott Education Consulting. You may share this summary freely in its complete, unaltered form. Not to be adapted, sold or reproduced in part without written permission.