CITATION MISREPRESENTATION IN THE CALIFORNIA MATH FRAMEWORK

Brian Conrad, Professor of Mathematics
Director of Undergraduate Studies in Math, Stanford University

EXECUTIVE SUMMARY

The Mathematics Framework Second Field Review (often called the California Mathematics Framework, or CMF) is a 900+ page document that is the outcome of an 11-month revision by a 5-person writing team supervised by a 20-person oversight team. As a hefty document with a large number of citations, the CMF gives the impression of being a well-researched and evidence-based proposal. Unfortunately, this impression is incorrect.

I read the entire CMF, as well as many of the papers cited within it. The CMF contains false or misleading descriptions of many citations from the literature in neuroscience, acceleration, de-tracking, assessments, and more. (I consulted with three experts in neuroscience about the papers in that field which seemed to be used in the CMF in a concerning way.) Often the original papers arrive at conclusions *opposite* those claimed in the CMF.

This Executive Summary includes highlighted samples of the misrepresentations of cited work within the CMF, followed by some brief comments on each of 7 areas of concern detailed in this document. These 7 areas (corresponding to sections of the main text) are:

(1) Neuroscience via Pseudoscience

(2) Devaluing of Advanced High School Math

(3) Myth from 1892

(4) Misrepresentations About Acceleration

(5) Misrepresentations About Tracking

(6) Misrepresentations About Assessment

(7) Miscellaneous Further Misrepresentations

Showing a citation is falsely described in the CMF does not mean a specific position is being advocated with respect to that citation. (There are other statements on content-based concerns with the CMF.) My purpose here is solely affirming that honesty and accuracy matter in public education policy. I found all misrepresentations and related inaccuracies when reading the CMF (and cited papers) on my own. I thoroughly checked everything in what I have written, but if there is an oversight or mistake then please tell me and I'll gladly fix it.

Examples of problematic citations in the literature, discussed in detail later, include:

(i) The CMF contains many misrepresentations of the literature on neuroscience, and statements betraying a lack of understanding of it. For example, the CMF claims that the "highest achieving people have more interconnected brains" and that creating "brain connections" improves learning. But the relation between brain connectivity and performance, especially in mathematics, is much more complex. In particular children with developmental dyscalculia exhibit abnormally *high hyperconnectivity* in several parts of their brain, contradicting the CMF's claims.

A sample misleading quote is "Park and Brannon (2013) found that when students worked with numbers and also saw the numbers as visual objects, brain communication was enhanced and student achievement increased." This single sentence contains multiple wrong statements: (1) they worked with *adults* and not students; (2) their experiments involved no brain imaging, and so could not demonstrate brain communication; (3) the paper does not claim that participants saw numbers as visual objects: their focus was on training the approximate number system.

Many more examples are discussed in §1. The nature of the errors implies (as explained in the main text) that whoever wrote these parts of the CMF lacks an understanding of the neuroscience literature regarding the learning of mathematics. False or misleading citation of papers cannot be used to justify public policy recommendations and guidance to districts. The neuroscience experts I talked with agree with a 2008 statement of the National Mathematics Advisory panel: "attempts to make these connections [of neuroscience] to the classroom are premature".

(ii) The CMF claims Ramani and Siegler (2008) showed that "after four 15-minute sessions of playing a game with a number line, differences in knowledge between students from low-income backgrounds and those from middle-income backgrounds were eliminated".

It would be great if eliminating educational gaps were as easy as playing a game several times. But Ramani and Siegler showed nothing of the sort. Their paper shows the non-surprising result that playing a game on a number line improves preschoolers' performance on a task of approximating the position of a number on the line. It does not say anything about differences in mathematical knowledge.

(iii) The CMF claims Liang et al (2013) and Domina et al (2015) demonstrated that "widespread acceleration led to significant declines in overall mathematics achievement." As discussed in §4, Liang et al actually shows that accelerated students did slightly better than non-accelerated ones in standardized tests. In Domina et al, the effect is 7% of a standard deviation (not "7%" in an absolute sense, merely 0.07 times a standard deviation, a very tiny effect). Such minor effects are often the result of other confounders, and are far below anything that could be considered "significant" in experimental work.

(iv) In yet another case, the CMF cites Burris et al (2006) for demonstrating "positive outcomes for achievement and longer-term academic success from keeping students in heterogenous groups focused on higher-level content through middle school". But the CMF never tells the reader that this paper studied the effect of teaching *Algebra I for all 8th grade students* (getting good outcomes) — precisely the uniform acceleration policy that the CMF argues against in the prior point.

(v) In some places, the CMF has no research-based evidence, as when it gives the advice "Do not include homework ... as any part of grading. Homework is one of the most inequitable practices of education." The research on homework is complex and mixed, and does not support such blanket statements.

(vi) The CMF claims that Sadler and Sonnert (2018) provides evidence in favor of delaying calculus to college, but the paper finds that taking calculus in high-school *improved* performance in college.

(vii) The CMF makes the dramatic claim that (Black et al, 2002) (really 2004) showed that students "are incredibly accurate at assessing their own understanding, and they do not over or under estimate it." If this claim were true, no exams would be needed to assess student knowledge: we could just ask them. Unfortunately, the paper of Black et al contains nothing of the sort.

Summary. In §1, I address many claims made using neuroscience that upon inspection are seen to be pseudoscience. In §2, I discuss citations that are misrepresented to promote the CMF's advocacy against the value of more advanced material of high school math for most students. In §3, I discuss the CMF's appeal to a false myth about a certain 1892 report that has repeatedly been stated in the media advocacy for data science, and show this myth to be false by going back to the original source. In §4, some citations that are claimed to show harmful effects of acceleration are seen to show nothing of the sort, and another citation about acceleration is shown to be misrepresented in a way that hides its supporting evidence in favor of acceleration.

In §5, some citations on tracking are shown to be misrepresented as supporting the CMF's advocacy for de-tracking (none of the cited references discussed herein provide such supporting evidence, instead giving evidence in the opposite direction or focused on an unrelated topic). In §6, many citations that are claimed to support the CMF's recommendations on assessment methods are shown not to do so. Finally, in §7 I collect the remaining misrepresented citations that I found.

Each label "Chapter x, lines y-z" in the main text discusses misrepresentations or false statements of citations in that part of the CMF. All documents cited here from the CMF are listed in Appendix B of the CMF, its overall bibliography (organized by chapter).

The abundance of false or misleading citations I found in the CMF calls into doubt the credibility of all appeals to the literature in the CMF. It is the responsibility of the California Department of Education to fix all defective citations. If there is neither time nor expertise to confirm the accuracy of a citation then it has to be removed, along with everything that depends on it.

Any further revision of the CMF needs to be done with a fresh set of people, given what has already transpired (as documented here, and in follow-up documents in preparation).

1. Neuroscience via Pseudoscience

The CMF appeals to many neuroscience papers to justify statements that are in no way supported by (and are sometimes entirely unrelated to the content of) the cited papers. I document such occurrences mainly for Chapter 1 because this issue arises in some other chapters (as will be identified in documents still being typed) but those are often repeating the problems found in Chapter 1. **I am not passing judgement on the cited papers, only on how they are invoked in the CMF.**

After I identified CMF neuroscience statements that seemed puzzling or questionable, I wrote up my concerns in detail and then consulted with several experts in neuroscience to refine the discussion of these problems. These experts are:

- David C. Geary (Dept. of Psychological Sciences, Interdisciplinary Neuroscience Program; University of Missouri-Columbia),
- Alan Jasanoff (Director of Center for Neurobiological Engineering, McGovern Institute for Brain Research; MIT),
- Allyson Rosen (Department of Psychiatry, Stanford University).

Just to reiterate: **I am not questioning the value or correctness of any neuroscience papers**. I am raising concerns about how such papers are invoked in the CMF for assertions far-removed from results in those papers.

1.1. **Chapter 1, lines 148-162**: This passage is full of highly inaccurate statements about neuroscience. The cited references as discussed below do not make conclusions consistent with the CMF claims (and often are far-removed from such claims). Therefore this passage has to be removed as matter of responsible professional education policy.

The CMF has the sentence:

> "Neuroscience research has shown that the highest achieving people have more interconnected brains – with different brain pathways communicating with each other (Menon, 2015; Kalb, 2017)."

This is false in multiple ways. Firstly, there has been work on the relationship between academic learning (relevant to the CMF) and brain development (such as (Shaw et al., 2006)) but the conclusions are not so straightforward to state as the CMF suggests. For example, (Holmboe, Fiske, 2019) says

> "Researchers have observed that structural pre-frontal cortex maturation consists of both progressive (myelination, neuron proliferation, synaptogenesis) and regressive (cell death, synaptic pruning, loss in grey matter) changes. Interestingly, the physical maturation of the frontal lobe appears to parallel the advances seen in cognitive abilities throughout childhood and adolescence."

Another flaw in the CMF's discussion here is that there is *nothing in the cited papers* that shows higher connectivity is correlated with higher math achievement. Firstly, since the

Kalb reference is a popular-science article (neuroscience in National Geographic by a professional writer), the CMF should not be citing it in lieu of research papers it may mention. Turning to the cited 2015 paper by Menon (which I will call Menon 2015a to distinguish it from a different 2015 paper of Menon and others cited later in Chapter 1), this is a review article that describes a specific group of functionally connected brain regions called the *salience network* that is argued to be important for explaining psychopathology (e.g., depression, Alzheimer's disease). There is no discussion of intelligence or achievement in this paper, and so no reason for the CMF to be referring to Menon 2015a.

Moreover, connectivity is not necessarily a positive attribute. For example, page 17 of Chapter 1 of the 2019 book *Cognitive Foundations for Improving Mathematical Learning* says:

> "The children with developmental dyscalculia exhibited abnormally high hyperconnectivity in frontal, parietal, temporal, and visual regions prior to training; in other words, they were engaging too many brain regions when attempting to order numerals."

So more connectivity is *not* always better: it's worse for kids with dyscalculia. Likewise, (Abreu-Mendoza et al., 2021) found evidence that

> "hyperconnectivity of the intraparietal sulcus and hippocampus, areas important for math cognition, remains a hallmark of low math ability into adolescence."

All discussion of "brain connectivity" in the CMF is pseudo-science and has to be removed. This is also in accordance with the findings of the Learning Processes task group within the 2008 National Mathematics Advisory Panel. That task group's detailed report noted the following on page 4-111 in section F (on Brain Sciences and Mathematics Learning) of its Reviews and Findings:

> "Brain sciences research has the potential to contribute to knowledge of mathematical learning and eventually educational practices, yet attempts to make these connections to the classroom are premature. Instructional programs in mathematics that claim to be based on brain sciences research remain to be validated."

Some progress has been made since the time of that 2008 recommendation, but in the view of experts in neuroscience the field is nowhere near ready for practice recommendations.

Now turning to the 2013 paper of Park and Brannon, the CMF says:

> "Park and Brannon (2013) found that when students worked with numbers and also saw the numbers as visual objects, brain communication was enhanced and student achievement increased."

Firstly, this paper is a study of *adults*, not students. Participants guessed the numbers of dots in large arrays of dots and then performed arithmetic. There was no brain imaging in this study at all, and it is entirely unclear what the CMF means by "brain communication". Furthermore, the paper has no claim that numbers are seen as "visual objects", and it isn't made clear precisely what such a statement even means. Park and Brannon suggest in the discussion at the end of their paper that "the non-symbolic approximate arithmetic task [the dot approximation task] actually reflects a sharpening in the Gaussian distributions for each numerosity". Here again, there is nothing about "student achievement" as the CMF vaguely suggests.

In fact, there is an entire literature associated with the type of work in the paper of Park and Brannon; and there is considerable debate (see (Qiu, et al., 2021)). Park and Brannon focused on the approximate number system (ANS), an evolved system for approximating the quantity of sets of objects; this engages part of the parietal cortex called the intraparietal sulcus and part of the prefrontal cortex. It is unclear whether acuity of the ANS contributes to formal math learning, and if it does then the effect is restricted.

Later in this paragraph of the CMF, the following is said about a 2008 paper of Ramani and Siegler:

> "Researchers even found that after four 15-minute sessions of playing a game with a number line, differences in knowledge between students from low-income backgrounds and those from middle-income backgrounds were eliminated."

This is outlandish: not only is "differences in knowledge" extraordinarily vague, but the paper is not about the substantial removal of anything that could be reasonably called "differences of knowledge". The paper tested preschoolers ability to order numbers between 1 to 10 on the line. Unsurprisingly, their performance improved after playing a numerical game in which the child moves a token on a number line marked with these numbers. Ramani and Sigler showed that low-income preschoolers performance on this particular skill improved to be comparable the baseline level of higher-income preschoolers, but did not test by how much does playing such a game improves the performance of the higher-income population.

The paper of Ramani and Siegler does *not* say anything at all about addressing significant differences in math knowledge among students from different socio-economic backgrounds. The dramatic suggestion by the CMF that four 15-minute sessions of a number line game can cause real "differences in knowledge" to be "eliminated" is absurd (if the CMF only intended to suggest a modest effect there would be no purpose to this being discussed in the CMF at all). The way the paper of Ramani and Siegler is being described in the CMF is pseudo-science, and so it has to be removed.

The Ramani-Siegler study and follow-up studies reveal gains in number knowledge for young children who have probably had low exposure at home. These are useful interventions for what they do, but do not support the CMF's dramatic claim quoted above. Moreover and more broadly, intervention effects generally fade out (i.e., students who did not receive intervention catch up); see (Bailey et al., 2020). There is a lot that remains to be learned about developing interventions with sustained impact.

1.2. **Chapter 1, lines 382-409**: This passage is full of highly inaccurate statements about neuroscience. The cited references (many of which I read) do not have conclusions anything at all like what is claimed, as explained below. Therefore this passage has to be removed as matter of responsible professional education policy.

The CMF says on lines 382-383 that

> "Another meaningful result from studies of the brain is the importance of brain connections."

Citing a 2015 paper by Menon et al. (which I'll call (Iuculano, 2015) because Iuculano is the first author named on the paper) that focused on brain activity when solving math problems, the CMF says

> "They found that even when people are engaged with a simple arithmetic question, five different areas of the brain are involved, two of which are visual pathways. The dorsal visual pathway is the main brain region for representing quantity. Menon and other neuroscientists also found that communication between the different brain areas enhances learning and performance."

This is a tremendous misrepresentation of what (Iuculano, 2015) is about, and is contradicted by other research (as I explain below).

The focus of the work (and its findings) in (Iuculano, 2015) is *not* about identifying which areas of the brain are involved in work on simple arithmetic questions. Indeed, the paper refers to earlier work for analysis of some areas that seem to be involved, and it is entirely about analyzing the effect of intensive one-on-one tutoring sessions for students with *math learning disabilities* (MLD). This work has *nothing* to do with "communication between the different brain areas"; it analyzes how activity levels of various brain regions for MLD kids changed in response to intensive tutoring.

There are many brain areas engaged when students solve arithmetic (or any other math) problems, including those noted in the CMF, but these areas become increasingly integrated with lots of practice, which results in their automatic engagement when presented with such problems (see (Qin, et al., 2014) for some research on the topic). Some areas drop out of the process (e.g., the hippocampus) and other areas become more engaged. Moreover, *any and all complex cognitive processes will involve communication between different brain areas*.

The actual main discovery of (Iuculano, 2015) is that in the short term, intensive tutoring on MLD children has the effect of making their brain activity during some simple arithmetic tasks become statistically indistinguishable from that of typically developing children. Moreover, like all good scientific papers, the conclusions are kept to within the limitations imposed by the scope of the experiments, noting near the end that generalization to more complex math problems and long-term effects would require further studies. A bit is known about the brain systems involved here, but nothing at a level that would support policy recommendations.

The preceding distortions and misrepresentations of the neuroscience literature have to be eliminated from all discussions informing educational policy related to math. There is no reason the CMF should be discussing the preceding matters, let alone appealing to the neuroscience literature (moreover in a false way) on the matter of learning and performance being enhanced by "communication between the different brain areas". After all, studying or homework also "enhances learning and performance", yet the CMF doesn't advocate for either of those. In fact, on lines 753-755 of Chapter 12 the CMF offers the guidance

> "*Do not include homework, in given, as any part of grading.* Homework is one of the most inequitable practices in education; its inclusion in grading adds stress to students and increases the chance of inequitable outcomes."

No research is mentioned to back up these claims. There is considerable debate on how much homework is appropriate (depending on a student's age and coursework), so the CMF should not be making such absolute statements.

Examples of research supporting positive aspects of age-appropriate amounts of homework include (Cooper et al., 2006) (which highlighted the need for more research) and some items mentioned within the more recent summary (Bempechat, 2018). Recent research is also revealing that suitable levels of stress can be an important aid in learning, so eliminating stress may sometimes harm students' abilities to learn; see (Rudland et al., 2020). This illustrates that the CMF is not only wrapping itself in a pseudo-scientific aura of neuroscience; it also cherry-picks and ignores very relevant results from neuroscience and related fields.

[It is also interesting to note that (Iuculano, 2015) uses multivariate statistical methodology including Support Vector Machines, a tool merging high-dimensional linear algebra and multivariable calculus. That is professional data science, out of reach of being learned by those who follow the CMF's later "advanced data science" guidance in Chapter 5 that avoids contact with Algebra II and is an off-ramp from the pathway to calculus.]

A similar distortion of findings occurs with the description of the cited reference to the 2013 paper of Park and Brannon immediately thereafter: the CMF says Park and Brannon found that (boldface mine):

> "different areas of the brain were involved when people worked with symbols, such as numerals, than when they worked with visual and spatial information, such as an array of dots. The researchers also found that **mathematics learning and performance were optimized when these two areas of the brain were communicating with each other**. Learning mathematical ideas comes not only through numbers, but also through words, visuals, models, algorithms, multiple representations, tables, and graphs; from moving and touching; and from other representations. But when learning reflects the use of two or more of these means and the different areas of the brain responsible for each communicate with each other, the learning experience improves."

There is *no basis* in the Park and Brannon paper for the CMF to draw conclusions about optimization of mathematical learning and performance, since that is not at all what the paper explores.

The paper reports on some basic experiments with visual and symbolic information processing, suggesting some ideas to explore for those who struggle with math; the dramatic way its conclusions are described in the CMF is a genuine distortion. The final CMF sentence above,

> "But when learning reflects the use of two or more of these means and the different areas of the brain responsible for each communicate with each other, the learning experience improves."

is not remotely justified by the cited references.

The CMF repeatedly refers to a relation between math and spatial abilities, but the specifics of this relation are far from being understood at a cognitive level, much less at the level of brain systems. A recent review of work in the area is (Ansari, Hawes, 2020).

The entire discussion about these neuroscience papers and all claimed conclusions from them have to be removed; it is pseudoscience.

1.3. **Chapter 2, lines 517-528**: This is pseudoscience and a misrepresentation of neuroscience literature, explained at length in my comments for lines 148-62 of Chapter 1, from within which this is a verbatim copy of all but its last few lines. Therefore it has to be removed.

2. Devaluing of Advanced High School Math

2.1. **Chapter 4, lines 146-159**: This passage on an ACT-organized survey of high school and college faculty is presented under the premise that instructors have to choose between depth of coverage and breadth of coverage in math classes. That premise is a false choice, as any experienced teacher of mathematics or any other subject knows well (these options are not mutually exclusive), but the real problem is that the CMF significantly misrepresents such survey results to support its preferred conclusions (as we shall now see), so it has to be removed.

It is noted that the survey indicates high school teachers generally put more value on the coverage of advanced topics and college instructors tend to put more value on mastery of the fundamentals. The CMF is setting this up as evidence for its preferred (false) narrative that it is best to cover more basic material more deeply, as that is then used to argue (again disregarding all nuance) that acceleration in high school is detrimental to student success in college-level math. Indeed, the CMF says on lines 157-158 in reference to the preference of high school teachers that:

> "This misunderstanding about the types of experiences that best prepare students for college mathematics success produces high-school graduates who enter college with a superficial grasp of superfluous procedures and little conceptual framework."

However, the CMF is grossly distorting what the ACT survey results actually say. To explain this, let's first quote lines 152-156 of the CMF:

> "[in 2012] almost all topics rated by college faculty as most important for incoming students are typically taught in grade nine or earlier (ACT, Inc., 2013, p. 6). Again in 2020, the top ten most important skills for incoming students, as rated by instructors of entry-level college math courses, are grade nine (or earlier) topics (ACT, Inc., 2020, p. 11)."

This sounds surprising, so let's look at the original ACT references cited to see what they say. The CMF passage turns out to be highly misleading, as we now explain.

Firstly, the CMF is focusing solely on Table 1 in the 2012 ACT survey results (on page 7) which is about the most important prerequisite skills for (credit-bearing) first college math courses – this is *not about calculus*, but rather about the broader notion of "first college math course". So this survey output is highly affected by the fact that the most common *first* college math course is Algebra II, aka College Algebra. (In fact, the middle of page 12 of the 2020 ACT document says explicitly that College Algebra is "the most common entry-level college math course". There is no reason to believe it was ever not this in prior years, such as 2012.)

Although UC and CSU campuses in the past gave no credit for college algebra, the ACT survey was not California-specific and college algebra does earn credit at many post-secondary institutions. Hence, for ACT's 2012 national survey one has to regard enrollment in College Algebra as having a huge impact on the results. Given that College Algebra corresponds to the math of grade 11 in high school, and grade 10 is geometry, it is *unsurprising* that the most important skills for success in this are learned in grade nine (i.e., Algebra I) or earlier.

Furthermore, if one reads *further* into the 2012 ACT document to look at Table 2 on page 9, which focuses specifically on the skills most important for success in precalculus and in calculus (rather than on the "first college math course"), one sees *lots* of topics (in fact, the vast majority) from later grades of high school math listed as crucial for such success. Again, this is no surprise, but also makes clear that the CMF's presentation of the 2012 ACT survey results is a misleading distortion.

It is interesting to note that page 7 of the 2012 ACT document emphasizes the critical importance of math classes revisiting earlier skills with new depth. That goes against the CMF's narrative that redundancy should be squeezed out wherever possible at the high school level (as part of the CMF's ideological desire to try to cram the usual 3 years of high school math into 2 years), but the CMF doesn't try to address that inconsistency (preferring to cite the ACT document only for aspects which support its preferred conclusions, even if some misrepresentation is required).

Turning to the 2020 ACT document, is the CMF's "top ten" claim true? No, it is false: items 5, 7, and 8 on the list come from grades 10, 10, and 11 respectively. But there is another issue hiding in plain sight, making the comparison with the 2012 survey apples and oranges: the 2012 survey focused on students' first college math course and the 2020 survey focused on "entry-level college math": the most introductory math taught at each college. That is a *very different* type of survey; e.g., at a college which offers College Algebra, that is the entry-level math course and so is the only one in the survey there. So the impact of College Algebra would be even more significant than in 2012, and once again the dominant role of topics from grade 9 and earlier is entirely to be expected.

The bottom line is that the ACT survey results from 2012 and 2020 are heavily misrepresented in the CMF both in terms of what they say and in terms of their (ir)relevance to readiness for calculus (for which the 2012 ACT survey gives exactly the type of results one would expect, going entirely against the narrative promoted here by the CMF).

This all has to be removed from the CMF, due to its heavy misrepresentation of the survey findings.

2.2. **Chapter 8, lines 1044-1047**: Here the CMF appeals to a paper (Sadler, Sonnert, 2018) as if that paper gives evidence in favor of delaying calculus to college. But the paper's message is *opposite* what the CMF is suggesting.

The paper controls for various things and finds that mastery of the fundamentals is a more important indicator of success in college calculus than is taking calculus in high school. There is nothing at all surprising about this: mastery of the fundamentals is most important. The paper is simply quantifying that effect (this is the CMF's "double the positive impact"), and also studying some other things.

What the paper does *not* find is that taking calculus first in college leads to greater success in that course. To the contrary, it finds that for students at *all* levels of ability who take calculus in high school and again in college (which the authors note near the end omits the population of strongest students who ace the AP and move on in college) do better in college calculus than those who didn't take it in high school (controlling for other factors). The benefit accrued is higher for those who took it in high school with weaker background, which again is hardly a surprise if one thinks about it (as Sadler and Sonnert note, that high school experience reinforces fundamental skills, etc.).

If one only looks at the paper's abstract then one might get a mistaken sense as conveyed in the CMF about the meaning of the paper's findings. But if one actually reads the paper, then the meaning of its conclusions becomes clearer, as described above. (Curiously, the paper's heavily statistical exposition involving F-tests and eigenvalues is an application of the tools of calculus and advanced algebra. There may be a lesson in that for those who denigrate these areas of math.)

3. MYTH FROM 1892

3.1. **Chapter 5, lines 1590-1593**: The myth here about the Committee of Ten from 1892 – that it promoted a high school math curriculum specifically focused on preparing for calculus – is false but has been repeated ad nauseam in the media for at least several years. It has to be removed.

If one goes back (as I did) to read the mathematics section on pp. 104-116 and the general subject-area grade-level recommendations for math on pp. 35-51 of the original

1892 report (which all CMF writers and CFCC members could have done since the entire report is linked near the bottom of the Wikipedia page about the Committee of Ten), one can see that the way this committee's work on math is described in the CMF is highly misleading and false.

The high school course sequence recommended by the committee was *not* specifically designed for calculus preparation. Indeed, the report explicitly included *two* options, one having "bookkeeping and commercial arithmetic" for sophomore and junior years as alternatives to further algebra. It also required geometry, and only those seeking scientific or technical degrees were recommended to take a 4th year of math, in trigonometry. This was a sensibly balanced proposal, not skewed toward the goal of calculus.

There is no mention of calculus anywhere in that report, and the fact that the pathway advocated for those planning to pursue a scientific or technical degree in college consisted of material in algebra, geometry, and functions leading to calculus is hardly a surprise: scientific and technical work uses exactly that mathematics (even if one doesn't need calculus). Bear in mind that continuing in school past 10th grade was uncommon in this country until around halfway through the 20th century.

Interestingly, the report of the Committee of Ten was also *anti-tracking*: insisting that everyone be at the same level in Math through the end of 9th grade, with a first Algebra course being the focus of 9th grade (after some basic symbolic manipulation exposure in 8th grade). So overall, this report from 1892 shares many similarities with the goals of the present CMF, making the denigration of that committee in the CMF ironic.

The US math curriculum underwent a dramatic change after Sputnik, updating the guidance from the 1892 committee precisely to meet the newer needs of those times. Have calculus and Algebra II now become less relevant? No: there are now many new sources of motivation, and the entirety of data science and machine learning rests critically on much of the conventional high school math content. Responsible and reliable use of computers to implement quantitative modeling requires familiarity with the foundational math content from algebra, geometry, and functions underlying those models.

The CMF has to stop spreading the false myth about the Committee of Ten from 1892*.* That myth has misled CFCC members into believing incorrectly that the content of the conventional math curriculum is obsolete. It has also misled others in positions of authority (such as staff in the UC Office of the President) into believing incorrectly that the traditional math content is "limiting" (see slides 13 and 17 of this slideshow), whereas in reality the traditional math content (which can certainly be provided with more contemporary motivation) keeps all STEM options open; it is not "limiting" at all.

I have only ever seen one discussion in the media for which the author clearly read the original 1892 report, and that author is also a mathematician.

4. Misrepresentations About Acceleration

4.1. **Chapter 8, lines 830-842**: This paragraph discusses two papers (Liang et al., 2012) and (Domina et al., 2015) on Algebra I in 8th or 9th grades. The conclusions of both of these papers are substantially distorted by the CMF, which refers to them when saying

> "several studies found that ... widespread acceleration led to significant declines in overall mathematics achievement".

We will see below that the two cited papers show nothing of the sort.

Let's first consider (Liang et al., 2012). In that paper, Table 1 shows near the left side that among those who took 8th grade Algebra I, the percentage proficient on the California Standards Test (CST) is 39% in the earliest cohort 2003-2006 and grows a bit to 42% for the last cohort studied (2008-2011); subtracting these from 100% is where the CMF's "approximately 60%" failure rate must come from. Among these who continue on to higher math, there are somewhat bigger gains over time in the passing percentage (see Table 1 reading down columns, or look in the lower part of the right column of page 333 in that paper). Note this is all prior to the 2013 Common Core.

Table 3 in that paper records the outcome of various math CST's (e.g., general 8th grade math, algebra, geometry, etc.) in 2007 for students who took the "general math" CST in 8th grade, and so the 2nd row of that table is those who took the CST for Algebra I for the 1st time in 9th grade. This is the comparison group within this study for whether Algebra I is "better" to be taken in 8th or 9th grades (and consists of more than 25,000 students). Let's see how the success rates compare. The far-right entry in the 2nd column of Table 3 shows this group of students was 82.86% of all 9th graders (note the entries in the right column sum to 100%), which the data partitions into failing and passing groups as 51.40% and 31.46% respectively (note that 51.40 + 31.46 = 82.46). Thus, the fraction of this non-accelerated subpopulation that was judged non-proficient is

$$51.40/82.86 = 0.62 = 62\%.$$

This is essentially the *same* as the non-proficient rate of the accelerated group that the CMF is complaining about among the accelerated group (and is actually a bit worse).

So the paper (Liang et al., 2012) demonstrates *nothing* at all like what the CMF purports it does. There's no doubt that readiness for 8th grade Algebra I is an issue with many complexities. But proposals such as all-in-8th grade (without a commitment of sufficient resources for readiness) or all-in-9th grade (which is a violation of the Math Placement Act, as well as a denial of reality) are both too simplistic. The CMF should take to heart a statement near the end of (Liang et al., 2012):

> "The statistical analyses presented here tell us what has happened relative to student performance in mathematics but cannot answer why the performance patterns emerge as we show them."

Turning to (Domina et al., 2015), the CMF claims it demonstrates "large negative effects" of 8th grade Algebra I on performance in the high school exit exam. Let's see what this "large" effect is by looking in the actual paper: it is 7% of a standard deviation. Not a standard deviation, nor half of a standard deviation, nor even a tenth of a standard deviation. The effect is 0.07 times a standard deviation (this is indicated in the text of the right column of the page 284); this is very far below the threshold for statistical significance in experimental research. Moreover, this data is all about the *mean* score, nothing else.

To summarize: a lot of kids enrolled in 8th grade Algebra I and there was a tiny negative effect on the *mean*. Calling that a "large negative effect" is very misleading. The CMF has to avoid such mischaracterizations and cherry-picked papers. The Discussion section at the end of (Domina et al., 2015) notes that other studies on the same topic reached different conclusions, and it tries to resolve the inconsistency. The CMF never discusses inconsistencies in the math education literature; that is irresponsible when citing papers as a basis for public policy, because then it becomes difficult to distinguish evidence-based guidance from ideology.

These distortions indicate an ideological (rather than evidence-based) opposition to acceleration. Providing advanced courses and opportunities in public schools are a matter of equity. Indeed, wealthier parents can and will send their children to private schools that provide these opportunities or hire tutors for their children. (The explosion of extra-curricular outlets such as Russian School of Math, Beast Academy, and so on that cover the conventional curriculum demonstrates this.) Students with such desire who come from lower-income families do not have these opportunities outside public schooling, so denying such options in the public schools puts them at a disadvantage later. This is one of the motivating rationales behind Adrian Mims' Calculus Project.

4.2. **Chapter 9, lines 179-181**: The paper (Burris, Heuburt, Levin, 2006) on a US-based study is cited for demonstrating

> "positive outcomes for achievement and longer-term academic success from keeping students in heterogeneous groups focused on higher-level content through middle school."

One naturally wonders: what is that higher-level content in middle schools? Looking at the paper, it is Algebra I for all 8th grade students (in a "diverse suburban school district"). The CMF omits to mention this extremely pertinent fact.

In other words, the study shows benefits from *accelerating all students to take Algebra I in 8th grade* and so the CMF cites the paper for its great success in heterogenous education but doesn't tell the reader that this was done with Algebra I in 8th grade (that wouldn't fit the CMF's preferred narrative). This omission puts the focus on heterogeneity and away from the highly relevant context of 8th grade Algebra I. I am not claiming Algebra I should be taught to everyone in 8th grade, but such an omission with that citation is a deception upon the reader.

The misleading nature of the citations implies that this entire passage needs to be removed or else the paper's main conclusion about acceleration needs to be presented accurately as above.

5. Misrepresentations about Tracking

5.1. **Chapter 9, lines 166-178**: Here the CMF promotes its anti-tracking narrative, and engages in numerous examples of misrepresented citations. This all has to be removed.

One meta-analysis (Rui, 2009) of 15 studies from both inside and outside the US is cited, indicating that

> "students taught in non-tracked groups that offer a more ambitious curriculum tend to have a higher achievement overall,"

noting that high-achievers aren't affected while low and middle achievers show "significant increases". But there are also plenty of studies which show that tracking has beneficial effects (some such will be noted shortly), so the issue isn't at all settled. Hence, the CMF's decision to cite only that which supports its preferred conclusion while not acknowledging that the issue remains a matter of reasonable debate (with the existence of ample evidence to the contrary) is unprofessional. The public deserves a more balanced discussion of pros and cons.

The review article (Woessmann, 2009) on tracking effects in many countries is claimed by the CMF to show that countries (many of which, apart from the US, have relatively homogeneous societies) that track students earlier than 9th grade "increase inequality in learning significantly". Let's set aside the problem that the CMF never defines its terms, so it isn't clear exactly what this quote even means. There is a more fundamental problem: this paper is *not about tracking as done in the US*. Indeed, on the first page it says:

> "In this article, tracking refers to the placement of students into different school types, hierarchically structures by performance."

Hence, the paper uses an *unrelated definition* of the word "tracking".

That tremendous difference in the meaning of the word "tracking" in the CMF (and in all discussion of that word in US education policy) versus in the cited paper explains another puzzle on the first page of the paper which otherwise would make no sense: the Table from 2004 on the paper's first page lists the US as beginning tracking at age 16. That is wrong if "tracking" is meant in the sense used in the CMF, but it makes a lot of sense under the definition actually given in the paper (since the vast majority of US students attend their local public schools, and so are not "tracked" in the sense of the paper).

Hence, the CMF's multiple appeals to (Woessmann, 2009) for the negative effects of early tracking are fundamental misrepresentations of the paper's definition of the central concept under consideration (tracking). Interestingly, the paper's Figure 3 compares each country's "education inequality" (precisely defined, unlike in the CMF: how the standard deviation of its score distribution on a certain international comparison reading test differs from the mean of such standard deviations across all countries in the study) in 4th grade and at age 15, showing that for the US this "inequality" *decreases* by a factor of 2 (from slightly above 10 to slightly above 5. This is actually a measurement which has real meaning because it has nothing to do with tracking, being a general comparison of international assessments. But it doesn't fit the CMF's narrative about educational inequality, so it is passed over in silence. (To be clear, even this latter measurement is irrelevant since it is about a reading test rather than a math test, but the CMF is choosing to cite this paper in the first place.)

5.2. **Chapter 9, lines 181-184**: Three papers cited here on the virtues of avoiding tracking are said to show that high-achieving students "are advantaged when they are given opportunities to extend work and discuss mathematical connections in non-tracked groups". But among these three papers, one of them studies an experimental group which was given better-trained teachers (making it hardly a surprise that everyone benefited), another takes place in Finland where the teachers have far superior professional development opportunities than in the US, and the third has been the subject of some controversy.

The misleading nature of the citations implies that this entire passage needs to be removed.

5.3. **Chapter 9, lines 287-299**: Here the CMF says it is discussing a "de-tracking initiative" in math in a suburban New York district for grades 6–8, and focuses on the heterogeneity of classes. The culminating class in grade 8 is described as

> "the first course in an integrated mathematics sequence incorporating algebra concepts (entitled Sequential Mathematics I) in eighth grade."

The CMF describes the effects in glowing terms:

> "The researchers found that the students who learned in heterogeneous classes took more advanced math, enjoyed math more, and passed the [statewide final exam] a year earlier than students in traditional tracks. Further, researchers showed that the advantages occurred across the achieve spectrum for low and high achieving students (Burris, Heubert, Levin, 2006)."

So here there are a variety of very positive outcomes, all attributed to the CMF's recommended methods "de-tracking" and "heterogeneous classes". But New York's Sequential Mathematics sequence was exactly an early incarnation of the Integrated Math rearrangement of content (I know this because I grew up in New York at the time it was first introduced), so here is a more direct description: it was an *acceleration* program to get everyone taking Integrated 1 in 8th grade. The tip-off that this is an acceleration program is that the cited 2006 paper of Burris et al. has "acceleration" in its title.

The CMF should not hide behind obfuscatory buzzwords and instead be crystal-clear that it is actually offering great praise for teaching the Grade 9 math class to everyone in 8th grade. (I am not personally advocating that everyone should take Algebra I in 8th grade. Rather, I am just noting the actual meaning of what the CMF is praising in a specific paper it chooses to cite.)

6. Misrepresentations About Assessment

6.1. **Chapter 12, lines 221-228**: This description of a 1998 study by Black and Wiliam on methods of assessment is a substantial misrepresentation, oversimplifying a complicated process, and hence it has to be removed.

To be precise, here the CMF makes the dramatic unqualified claim that:

> "if teachers shifted their practices and used predominantly formative assessment, it would raise the achievement of a country, as measured in international studies, from the middle of the pack to a place in the top five."

(Here, *formative* assessment is essentially measuring knowledge during the learning process to provide instant feedback, as opposed the more traditional cumulative assessment that is called *summative*.)

These claims are so hard to believe that one has to look up the study to see if this is really what it says. The answer is negative. Firstly, the study acknowledges that teachers develop effective formative assessment *slowly*, via professional development. It is not any kind of "quick fix", as the CMF seems to be suggesting here. Also, Black and Wiliam were analyzing the effectiveness of formative assessment *along with* an array of other innovating teaching practices, about which the CMF says nothing. Furthermore, the claim about rising to the "top five" is the *maximum* effect that was seen, not at all representative.

The most damaging flaw in the CMF's oversimplified (and inaccurate) description of the conclusions of Black and Wiliam is that the CMF gives essentially no guidance on *how* teachers should become proficient at formative assessment whereas (Black, Wiliam, 1998) notes the cautionary warning:

> "Teachers will not take up ideas that sound attractive, no matter how extensive the research base, if the ideas are presented as general principles that leave the task of translating them into everyday practice entirely up to the teachers. Their classroom lives are too busy and too fragile for all but an outstanding few to undertake such work."

Finally, the CMF refers to both the 1998 paper of Black and Wiliam, as well as a follow-up, by Black et al, for the striking claim that:

> "if teachers were to assess students formatively, then the positive impact would outweigh that of other educational initiatives, such as reductions in class size."

I read both references from front to back, and found *nothing* in them to support any statement of this type. Where *exactly* do Black and Wiliam give evidence for such a claim within the cited references?

[I do not understand the absence of exact page numbers in most CMF citations, The entire purpose of a citation is so a reader can look it up and read the evidence. By omitting exact page numbers, a reader is left on a wild goose chase. In mathematics research papers, the tradition is to generally give *pinpoint-precise* exact references *within* papers or books so a reader can go straight to the relevant part. The CMF has to systematically tell the reader on exactly what page(s) to look in cited documents.]

Chapter 12, line 599: Here the CMF considers the choice among three options for evaluating classwork: giving it a grade, giving diagnostic feedback and no grade, or giving both such feedback and a grade. Work on this topic in (Butler, 1987, 1988) is cited, and it is said that the cited research shows groups of 5th and 6th grade students who got feedback and no grade "achieved at significantly higher levels" than such groups that got either of the other two treatments (for which the group-level achievements were comparable). We shall see that the CMF significantly misrepresents the scale and scope of Butler's work, so this all has to be removed.

It's unclear what "achieved at significantly higher levels" means, especially when comparing with a group that seems to have never received numerical grades. We will come back to this. The CMF also says Butler arrived at some conclusions for the top and bottom quartiles by GPA within each experimental group:

> ". . . both high-achieving (the top 25-percent grade point average) and low-achieving (the bottom 25-percent grade-point average) fifth and sixth graders suffered deficits in performance and motivation in both graded conditions, compared with students who received only diagnostic comments."

But the CMF doesn't explain how a group that gets only diagnostic feedback and no grade has any meaningful concept of GPA (= grade-point-average), a puzzle which will be demystified when we discuss what Butler *actually* did (which is not what the CMF writing suggests).

The papers by Butler are not listed in the CMF bibliography, but I was able to determine the identity of the later paper (Butler, 1988). Setting aside the vagueness of the CMF's writing, the more serious problem is that the CMF's description of Butler's work is a significant misrepresentation. This continues a long tradition, since (Guskey, 2019) points out that for those who advocate in favor of diagnostic feedback and no grades,

> ". . . writers and consultants typically cite a study conducted by Ruth Butler from Hebrew University of Jerusalem in 1988. But it seems clear that few have read the original study and understand its focus or the nuances of its findings. Instead, they cite other authors' summaries and conclusions, without careful attention to these crucial details."

That paper nicely summarizes what Butler actually did, but we can also see the deviations from the CMF narrative by looking directly in Butler's paper, as follows.

Firstly, Butler's study involved a very basic language task (e.g., making words from a given small set of letters) and a couple of puzzle-type free-range questions, for a random sample of 132 students students outside the context of their actual classes. Hence, the distinction of grade or no grade was on these isolated tasks, which is nothing at all like running actual classes for many months with or without grades. The very limited setting for this study is never mentioned in the CMF, but (like any good scientific paper) the end of Butler's paper recognizes these serious limitations of the experimental design, where she writes as the final sentence:

> "However, further research is clearly necessary to clarify the effects over time of systematically providing such feedback in applied settings and of reducing the use of normative grades."

The fact (never mentioned in the CMF) that Butler's work focused only on a basic language task and a couple of open-ended puzzles raises genuine doubts that this has relevance to running math classes.

What about the impression created by the CMF presentation that Butler noticed effects only in the top and bottom quartiles? Butler writes in the top paragraph on page 4 of her paper that she *only* analyzed data for students in the top and bottom quartiles (of GPA as measured by the class they were actually in: this is GPA determined by "grades" having *nothing at all* to do with the grading process explored in the study). Hence, she has no effects to report for the middle half since she never analyzed their data; in the end, $(1/2)132 = 64$ students were analyzed. In fact, even fewer: from within each of the top and bottom quartiles, Butler randomly chose 22 students to analyze (for a total of 44).

So this is ultimately a study of just 44 students, and the CMF's description of Butler's measurements in terms of the feedback-only group having "achieved at significantly higher levels" (compared with the other two groups) is referring to *merely a basic language task and some mental puzzles*. Hence, such "achievement" has nothing to do with learning. Likewise, the study measured student interest only in the assigned artificial tasks; it is an exaggeration to claim as the CMF does that Butler's work tells us about effects on student motivation for learning in real classroom settings.

The CMF is stretching Butler's conclusions far out of proportion to their limited setting, and on the basis of this is making sweeping policy guidance for how to run actual math classes. Also, Butler is measuring statistically whether differences between experimental groups are "significant" *in the sense of statistics*: is some difference in effects very likely to be genuine (not a statistical accident)? That is *not at all* related to the CMF's use of the phrase "significantly higher" in reference to differences between achievement levels in Butler's work, so the CMF's wording is a misrepresentation (or misunderstanding) that the use of the word "significant" in statistics does *not* mean what it does in ordinary speech.

Overall, the CMF is significantly (in the usual sense of the word) misrepresenting the scope and the relevance of Butler's study to the setting of state-level policy guidance for math classes.

6.2. **Chapter 12, line 610**: Continuing to promote its ideological stance against grades in math classes, the CMF claims that the paper (Pulfrey, Buchs, Butera, 2011) (which does not appear in the CMF bibliography, but I could determine uniquely from the author list and the year) is a follow-up to Butler's work and shows:

> "... that students who received grades, as well as students who received grades and comments, both underperformed and developed less motivation than students who received only comments. They also found that students needed only to *think* they were working for a grade to lose motivation, resulting in lower levels of achievement."

This description from the CMF is false in numerous ways: the paper implies nothing about the connection between grading in US K-12 math classes and motivation towards learning because it has nothing to do with K-12 students, nothing to do with math, and nothing to do with motivation towards learning. In particular, this entire passage has to be excised from the CMF.

This paper is about 3 experiments with *college-age* students in Switzerland in an English-as-foreign-language class taking a single English comprehension quiz (under the various conditions of being told it will be graded and count towards the GPA, or be graded along with comments, or only be given comments). The paper's main conclusion is that having a grade incentivizes a shift in student goals: to avoid getting a low grade ("performance-avoidance") rather than doing better than classmates ("performance-approach").

This has no relevance to what the CMF is discussing here. Furthermore, in addition to merely being for a single quiz, each experiment involves just a bit over 100 subjects, which is far too small for trying to infer lessons for education policy. In the cited paper, the authors (like all good researchers doing small-scale experimental work) acknowledge several limitations of the study. The CMF disregards all such caveats when citing this work.

6.3. **Chapter 12, lines 643-644**: The paper (Black et al, 2002) (really from 2004) is cited in seemingly yet another false way (so once again this has to be removed): the CMF makes the dramatic claim that

> "when students are asked to rate their understanding of their work through self-assessment, they are incredibly accurate at assessing their own understanding, and they do not over- or underestimate it."

This claim is so extraordinary (even setting aside that it is extremely imprecise) that it is impossible to believe. So I looked up the cited paper, and I found *nothing whatsoever* in it to support this claim. Where is the real evidence for this unbelievable claim?

6.4. **Chapter 12, lines 729-756**: This "Advice on Grading" seems to consist entirely of the opinions on grading from a single CMF writer's own book and put directly into state policy.

These opinions are not supported by any cited evidence, and are contrary to the experience of those who have taught mathematics (or any STEM subject). It is inexcusable for essentially an opinion piece by a CMF writer to be inserted into the CMF. Based on these many reasons, this has to be entirely excised from the CMF.

Let's go through the "advice" items to see why they are problematic as state-level guidance:

(i) Item 1 says "Always allow students to resubmit any work or test for a higher grade". Firstly, teachers are already swamped for time; now the CMF recommends all work can always be resubmitted, thereby potentially increasing the grading work of teachers by a huge amount. This is a recipe for burnout. At a time when California is experiencing a crisis-level shortage of credentialed math teachers and an exodus from the teaching profession in California math classrooms, it seems irresponsible. Moreover, in a math class (unlike an essay-based class) there is the realistic possibility of kids copying from a solution set or from work of classmates for resubmission.

I am not saying that revise-and-resubmit cannot work in a math class, but rather that such a method has special challenges in math compared to many other subjects. Hence, realistic actionable guidance on this has to involve a lot more details and reliable citations of past studies.

As a blanket statement with no guidance on sustainable implementation, this is chaotic and disrespectful of the time burden on math teachers.

(ii) Item 3 says "Use multidimensional grading". All teachers know about the idea of a modest class participation grade and other little inducements to encourage kids to persist in their learning. But what else is being suggested here? Is the teacher supposed to be keeping track of a large number of new things for all students? Public school teachers have only so much time, and often have to do all grading on their own (unlike university faculty).

Where is the evidence that "multidimensional grading" (however it is precisely defined) is realistic and sustainable? Furthermore, the CMF advocates formative rather than summative assessment; how is that compatible with multidimensional grading for a large class?

This is all much too vague, with little in the way of actionable precise guidance. Is the teacher supposed to buy the CMF writer's book from which this is excerpted to find out more details?

(iii) Item 4 says that grading on a 100-point scale is "mathematically egregious" (what does that mean?), and says there should be just grades of 0, 1, 2, 3, 4. For all exams, even longer ones? With such a blunt system, how are kids who get 3 supposed to know whether they were close to a 4? When taking an exam, there is genuine information conveyed to a student when they score 82% versus 89%; now it should all be subsumed under "3"? How is there any notion of partial credit with this suggestion?

(iv) Item 5 opposes allowing a course grade to incorporate performance on early material which often contains a lot of review of topics from a prior course. That is ill-advised. Reinforcement of prior material has genuine value, and the opportunity to get a grade can also provide *incentive* for a kid to be attentive and put in an effort to learn the material better on a second pass. This is important because (unlike many other subjects) mathematics is *cumulative*.

If kids know a portion of the class will not count at all towards the grade then they are more likely to tune it out, and this can leave them with a shaky foundation upon which a lot of subsequent content depends. Furthermore, different kids may have forgotten or never understood different parts of a prior course, so disincentivizing attentiveness from the start can *reinforce* differences in preparedness rather than enable early review to bring more kids to a common level.

(v) Item 6 is the following surprising guidance:

> "*Do not include homework, if given, as any part of grading.* Homework is one of the most inequitable practices in education; its inclusion adds stress to students and increases the chances of inequitable outcomes."

Such a statement, moreover not backed up by research, has no place in the CMF. Firstly, there is considerable debate about the appropriate amount of homework (depending on age and coursework). Also, there are examples of research supporting positive aspects of of age-appropriate amounts of homework, such as (Cooper et al., 2006) (which highlighted the need for more research) and some items mentioned within the more recent summary (Bempechat, 2018).

There are also many real-life experiences that call into question the absolutist position in the CMF. Has anyone ever learned to play the piano without sustained practice? Or learned a foreign language without practice at home? Or play a sport well without practice outside school hours? And why disincentivize effort on homework by not counting it at all in the course grade?

Would any CMF writer or CFCC member be willing to fly in a plane designed by engineers who never did physics homework? Or use a dentist who never practiced in the clinic during dental school? Or hire a financial advisor who never did their math homework? Or take out loans for their child to attend medical or law school but urge them to never do homework while there?

It does the policy discussion no favors to phrase this in terms of absolutes. A balanced approach is more useful to teachers and school districts.

Calling homework "one of the most inequitable practices in education" contradicts the experience of anyone who has ever learned challenging material in an educational setting. What is truly inequitable is to *not* give homework, and to *actively disincentivize* kids from getting some reinforcement and practice by refusing to incorporate homework in a course grade.

6.5. **Chapter 12, lines 914-916**: The CMF cites (Brookhart et al., 2016) to justify the claim that:

> "parents are supportive of mastery-based grading, as an alternative to traditional grading."

Mastery-based grading (MBG) entails assessing students on a *binary scale* (mastered or not mastered) for each specific learning target in a list, without the use of numerically-averaged scores on homework or exams. The sample report cards on pp. 46-49 of Chapter 12 of the CMF demonstrate the binary intent. The discussion of MBG occupies lines 757-937 of Chapter 12.

But the cited paper is *not about MBG*. It is focused on an entirely different approach called *standards-based grading* (SBG) that is based on *progress* in relation to grade-level standards. Near the top of the right column on page 22 of the cited paper it is explicitly noted that SBG is different from MBG. The CMF says on lines 705-708 of Chapter 12 that

MBG "is sometimes referred to as standards-based grading" but this is misleading: the two concepts really are not at all the same. (Some confusion may arise because what is now called "standards-based education" used to be called "mastery learning"; see page 2 of (Brookhart et al., 2016).)

The procedure laid out in the CMF is definitely MBG and *not* SBG. Indeed, in addition to following the binary scale, the learning targets in the CMF discussion are so vague and imprecise that they reduce mathematics to performance art. For example, the CMF's sample of learning targets for MBG is largely decoupled from any specific procedural fluency or *the ability to synthesize anything*. This is shown in the Table on line 890 of Chapter 12, which assesses multiple topics (e.g., piecewise functions, logarithmic functions, sequences and series) under the vague criterion "I can use, create, describe, and analyze [fill in topic] using different representations" (what happened to being able to use a math concept to solve specific types of problems or carry out specific tasks?).

The last three learning targets in that same Table are more distant from specific skills:

- I can demonstrate 8 mathematical standards.
- I can participate and engage in class/group discussion and problem-solving synchronously and asynchronously.
- I can take ownership over my own learning and develop positive identity as a thinker and learning of mathematics through reflection, self-determination, and grit.

Certainly allowing class participation and effort to contribute to a course grade is fine, but the above three as *learning targets* (setting aside that it isn't even clear what "demonstrate" in the first one means)?

The CMF confuses the reader into thinking SBG and MBG are the same, and it cites a paper on SBG to support its claims about the quite different MBG. Such confusion can make a paper on SBG appear to be relevant when it actually isn't. Also, the cited article's evidence about parental support for a given grading scheme is (like the entire paper) centered around SBG rather than MBG: this evidence comes up on lines 5-9 in the left column of page 25 of the cited paper (fleshed out further in two entries for Table 6 on pages 23-24). So the actual citation of (Brookhart, 2016) in the CMF is a serious misrepresentation.

Finally, in the right column of page 22 of (Brookhart, 2016) it is noted with *six citations* that school districts and teachers "have experienced difficulties in implementing SBG", and so surely there are also such challenges for implementing MBG given that MBG entails binary judgements and vague learning targets. But the CMF passes over in silence any challenges for implementing what it proposes, which is unacceptable in a document whose purpose is to provide guidance.

To be clear: I am *not* passing judgement on mastery-based grading as a pedagogical approach. Rather, I am pointing out serious errors in how the CMF cites references for it and a significant inadequacy in the details of its guidance on MBG (to be useful to teachers and to avoid a lowering the knowledge level defined by the content standards).

7. Miscellaneous Further Misrepresentations

7.1. **Chapter 2, lines 973-981**: The description here of the article (Esmonde, Caswell, 2010) is a mixture of false and misleading, so it has to be removed.

Consider the "number book project" for a kindergarten class in this article that the CMF highlights. The purpose was for the children's families to share some stories. But the CMF omits the relevant fact that for the experience discussed in the article, there were *no responses from the families* (see page 11 of the article). How can the CMF promote an illustrative example from the literature and not tell teachers that in practice it was a dud? In particular, the final sentence here in the CMF (on lines 980-981)

> "They then design classroom activities that draw on these number stories, songs, and games."

never actually happened in the case where it was tried. Given the CMF's role as a source of guidance to math teachers, publishers, and school districts, it is a significant misrepresentation of a citation to present an idealized scenario and never mention that when tried it badly failed.

[As a more minor matter, the "water project" for a 5th-grade class highlighted here in the CMF is presented on pages 9-10 of the article (and referred to in some later places), but it has virtually *no math content* (and certainly nothing at the 5th grade level). Hence, it makes no sense for this to be discussed in the CMF.]

7.2. **Chapter 5, lines 295-296**: This quotation from Arnold (2007) is a fabrication: there is *no such statement* in that paper, and according to Google search this quote does not exist anywhere in the world except in this Chapter 5 and the May 2021 public comment on an earlier version of Chapter 5. So it seems that the CMF copied a quotation from earlier public comment without checking the cited source for accuracy. Either way, this has to be removed.

7.3. **Chapter 5, lines 1415-1422**: The example from the literature in Figure 5.11 is significantly misrepresented here and so has to be removed and replaced with something else. Indeed, if one looks up the original source (as I did), one sees that the plots represent NO_2 concentration as a *weighted average* (see the formula for C_j on page 3 of the cited 2017 paper of Clark) among 210,000 blocks of population organized by 1% ranges of nonwhite population in each. In other words, this is *not* representing change over time and space as said in the CMF, but rather over time and race (so to speak) where the effect of "space" is wiped out by the weighted averaging over population blocks.

So this is not in any way an example of "change over time and change in space". Hence, this has to be replaced with an actual example of such in order to illustrate the CMF's intended message about variation across time and space.

7.4. **Chapter 8, lines 1071-1084**: Here the CMF claims to be listing student competencies from a specific UC & CSU document. But that is false; what is actually listed are *not* student competencies. Rather, the CMF is listing advice from pages 4-6 of the same document for recommended *aspects of teaching* to support student understanding; that is a very different thing. Moreover, the first item listed (on line 1077) is "Modeling" but the actual item listed in the document cited is "Modeling Mathematical Thinking", which has a rather different meaning. This change of the first item fits the CMF's bias toward data science.

Rather than import a list from elsewhere with a mischaracterization of what it represents (and a first item transcribed incorrectly in a way which fits a known biased narrative of the CMF), the CMF should instead refer to a detailed list of student mathematical competencies within itself: the section "Key Mathematical Ideas . . . " that is presently buried out of sight at the end of Appendix A (and which I co-authored with two other mathematicians based on input from Stanford faculty across all STEM fields, and submitted to the prior CMF revision). That would be more convenient for the reader, and is similar in spirit to the actual student competencies listed in the cited UC & CSU document.

7.5. **Chapter 12, lines 59-63**: Here the CMF mentions that tests on purely procedural knowledge, involving no reasoning or problem-solving, are "of limited use in predicting success in college and the workplace" (hardly a surprise, and also of no relevance since the purpose of tests in K-12 class is to measure mastery of course material and not to predict college or workplace success). It then claims that this fact

> "has led leading employers, such as Google, to eliminate standardized tests from their application requirements (Bryant, 2013)."

That statement about what leading employers (and in particular Google) have done is false: if one looks up the reference (a New York Times article), it says that (standardized) tests and GPA are used (only) for recent college graduates. So the reality as far as students emerging from the educational system are concerned is opposite what the CMF is claiming here. Hence, this all has to be removed.

Chapter 12, lines 80-83: This passage is the false description of (Park, Brannon, 2013) that is discussed in the comments on lines 148-162 of Chapter 1. It therefore has to be removed.