Review of the 4th and 8th grade algebra and functions items on NAEP

By Hyman Bass
Brookings Institution September 15, 2005

Guiding Questions

About the NAEP Algebra Items (4th and 8th Grades), we have been asked to address the following questions:

1. Are the items mathematically sound?
2. Are the items assessing students' knowledge of algebra? Or something else (pre-algebra or algebra readiness)?
3. If students do well on these items, would that be strong evidence that they know algebra?
4. Are important algebra topics and skills not covered by these items?
5. Please single out one or two items that you feel are exemplary and one or two that you feel are deficient.

I. Mathematical Soundness

I take *mathematical soundness* of *an item* to mean that*:*

(a) Is clearly and precisely formulated, at a level accessible to the intended grade, and free of non-purposeful ambiguity.
(b) Treats substantial mathematical knowledge and/or skills, with due regard for the need of a wide range of levels of challenge in the ensemble of items.
(c) If contextualized, has a sensible context, one that is taken seriously, and one that does not obscure the intended mathematical focus of the item.

Further, mathematical soundness requires that:

(d) The *ensemble of items* mathematically represents, in a reasonably proportioned and comprehensive way, the mathematical landscape of the domain (in this case, Algebra) that it putatively covers.

I find that, with only minor exceptions[1], the NAEP items meet conditions (a) and (b). Condition (c) also seems to be generally met, with a couple of notable exceptions, such as the following:

> G8, #21:
> The context of this integer subtraction problem posits that the lowest point on the St. Lawrence River is 294 feet below sea level. I can imagine that many test takers will be as startled as I was at this implausible sounding claim, for which I have found no confirmation after a modest search; perhaps they are referring to the bed of the river. In any case, this could easily distract test takers from the simple mathematical focus of the item.

---

[1] For example, in G8:
#20: The polygons should be convex.
#38: "should be" should be something like, "must be" or "is"
#39: "what to do" should be something like, "what one could do"
In G4:
#16: "a rule used" should be something like, "a rule that could be used"

G8, #24 (= G4, #12):
A table is given showing the weight w = w(m) of a puppy after m months, starting with w(1) = 10, and giving w(m) for m = 1,2,3,4. It asks for w(5). Inspection of successive increases shows that w(m+1) – w(m) = 6-m, whence w(5) = w(4) + 2. So far, so good. But, if this model persists, the puppy starts to <u>lose</u> weight after 6 months, and, after 13 months, passes into anti-matter!

II. Validity

Validity of the ensemble of items signifies satisfactory answers to Guiding Questions #2, #3, and #4 above. Guiding Questions #2 and #4 are ostensibly mathematical, and correspond to condition (d) above:

(d) The *ensemble of items* mathematically represents, in a reasonably proportioned and comprehensive way, the mathematical landscape of the domain (in this case, Algebra) that it putatively covers.

Guiding Question #3, on the other hand, is psychometric as well as mathematical: We would like to know that test takers do not get answers right based on faulty thinking, or thinking unrelated to what we think we are testing for, and, conversely, that proficient thinkers are not led to wrong answers (due to faults in item construction). We on this panel are not in a position to make such judgments. One method would be to use cognitive interviews with test takers, to learn what they are thinking when they construct their responses.

Here is an illustration of one type of difficulty: Item G4, #14 presents the first few powers of 2, 2^n, n = 2,3,4,5, and asks, if this sequence were to be continued, whether the number 375 might appear as a value. The presumed intention of the item is to test for knowledge about even numbers, and the recognition that the odd number 375 could not appear among the even numbers 2^n. On the other hand, the student might well continue multiplying by 2, to get 2^n for n = 6,7,8,9, with $2^8 = 256$ and $2^9 = 512$. Since the numbers visibly increase, passing and missing 375, 375 cannot occur. This perfectly correct mathematical reasoning, leading to the correct answer, implies no knowledge of even numbers, and demonstrates a computational and investigatory skill perhaps not considered particularly algebraic.

While condition (d) would appear to be a mathematical one, it is in fact not entirely so. For what we mean by "algebra," as a school subject, has undergone fundamental transformation in recent history, and remains a matter of some debate. The two natural sources of a definition – the school curriculum, and the discipline of mathematics – provide rather distinct maps and portraits of the territory.

Very broadly speaking, the traditional treatment of algebra was as the theory of algebraic (= polynomial) equations. The complexity of the theory depends on three parameters: the number of variables, the number of equations, and the degrees. Degree one is linear algebra, and school algebra generally treats at most two equations in at most three variables. One linear equation in two variables is essentially the study of linear functions, and this is a central topic in all curricula. The traditional treatment gave much more attention to fluency with symbolic manipulation, powers and radicals, and included (in high school) the quadratic formula and some general observations about polynomials of high degree, for example the binomial theorem, and factorization of $x^n - y^n$.

The contemporary curricular treatments of algebra foreground the idea of functions and patterns. The underlying philosophy seems to be something to the effect that we live in a data driven world, and what mathematics pre-eminently contributes is the mathematical modeling of data. So functions typically appear not as closed formal expressions, but rather as discovered ways to encode a pattern, to be discovered in a numerical sequence or table. While this has a certain surface appeal, and supports a lot of interesting student problems, it is at root a misleading caricature of the modeling process to which it pretends to pay homage.

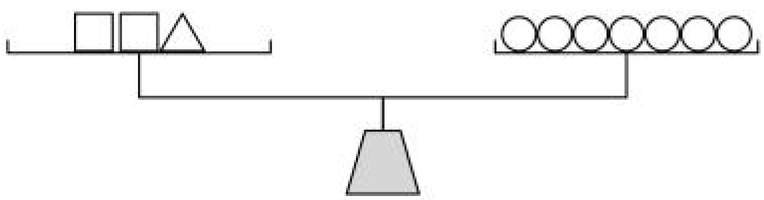
The world is complex and messy, and modeling real data requires a rich repertoire of functions with which to do anything serious. School students are today given a repertoire consisting only of linear and exponential functions, with perhaps a few quadratics, and so they are presented with data almost always of linear or exponential type. This is interesting material, but, to my mind, such pattern recognition occupies too much of the curricular space, and this is reflected in the NAEP problem sets.

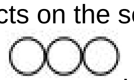
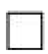
Among the things that one might imagine being included in a treatment of early algebra are various algebraic identities, sums of powers (geometric series), rules of exponentiation, simplification of numerical expressions involving square roots, primes and factorization, …. I am not necessarily arguing for inclusion of everything in this list.

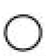
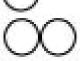
Rather than dwell on such debatable generalities, I tried to make a more fine grained, and non-ideological analysis of the mathematical features of the NAEP items we were given. In order to organize this, I had to construct categories of features that, collectively, could encode a reasonably faithful characterization of each of the items, along the mathematical dimensions I had identified. The results of this analysis are reported in the attached Appendix A. Appendix B is an annotation of the tables appended to the sets of NAEP items we were given.

III. Some exemplary item features

G4, #2

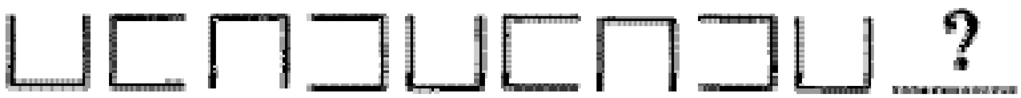


2. The objects on the scale above make it balance exactly. According to this scale, if [triangle] balances , then  balances which of the following?

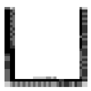
   A) 
   B) 
   C) 
   D) 

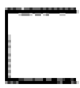
This models two linear equations (as balance beam) in three (iconic) variables. The student is asked to express one in terms of one other. This is a very attractive and substantial algebra problem accessible at 4th grade level. Note the design of the balance so as not to violate the physical law of the lever, needed for this to represent an equation. This is exemplary attention to features of the context.

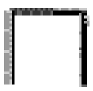
G4, #10

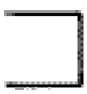


10. In the pattern above, which figure would be next?

A) 

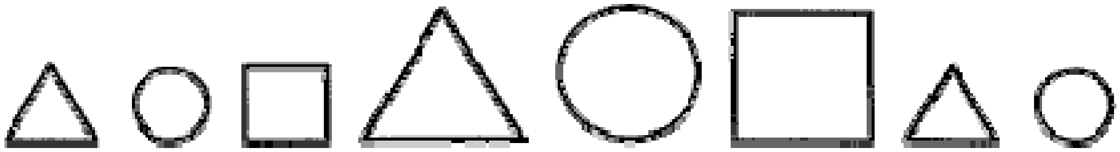
B) 

C) 

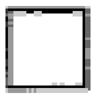
D) 

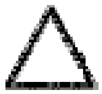
This geometric pattern extension problem gives an early glimpse of group theoretic ideas, looking and the orbit of a figure under the group generated by a -90^{o} rotation, leading to a period 4 pattern.

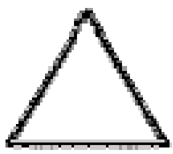
G4, #20

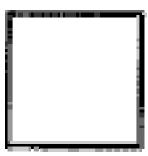


20. A pattern of shapes is to be repeated many times. The figure above shows one completed pattern and the beginning of the next. What shape comes next?

A) 

B) 

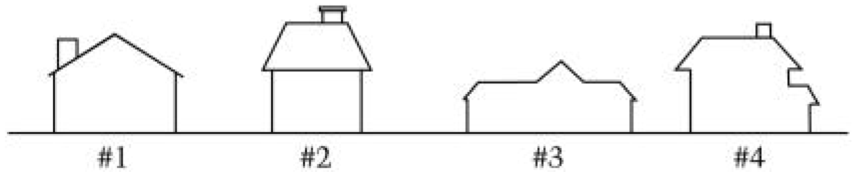
C) 

D) 

Did you use the calculator on this question?

O Yes O No

This geometric periodic pattern has a two dimensional variable figure, depending on 3 shapes and 2 sizes. I'm not sure how a calculator could be used here.



17. Allen, Bridgitte, Chaz, and Diann each live in a different house on the same side of a street. The houses and their numbers are shown above.

- Only one of the other three people lives next to Bridgitte.
- Chaz lives next to Bridgitte and next to Diann.

Which person could live in house number 2?

A) Allen only
B) Chaz only
C) Diann only
D) Chaz or Diann
E) Any of these four people could live in house number 2.

This is an attractively simple and accessible problem in combinatorial logic (about finite ordered sets and permutations). It calls on some nice problem solving reasoning.