Chapter 12
Procedural Steps, Conceptual Steps, and Critical Discernments: A Necessary Evolution of School Mathematics in the Information Age

Martina Metz and Brent Davis

Introduction

Early in 2010, the Organization for Economic Cooperation and Development (OECD) published an online document in which it distinguished among "formal," "informal," and "non-formal" education. Many elements of this typography were predictable. Of the three, for example, only formal learning was identified to involve certified teachers, accredited curricula, and institutionalized settings. But there were also some unexpected elements. In particular, not many educational leaders expected a prominent—but unexplained and unjustified—statistic asserting that 75–85% of one's learning is other than formal.

That sort of datum is hard to contest. In fact, it would seem reasonable to argue that it is grossly underestimated. While not fully explained in the report, one can infer that the number to indicate the portion of an average life not dominated by attending school. That is, it was intended to emphasize the importance of lifelong learning. If that was the purpose, the point is simultaneously important and trivial. And that is perhaps why some in the educational establishment viewed the statistic with suspicion, as a not-so-veiled move to diminish schooling's long-held authority in matters of defining, offering, and certifying learnings.

In this regard, the technological context of the OECD's pronouncement is significant. It was a statement on learning in the Information Age. With advancements in and ubiquity of communication and storage technologies, traditional schools can no longer maintain a pretense of guardians of and gatekeepers to cultural knowledge. While broad awareness of that pretense has not yet contributed to substantial transformation in the institution, it would seem reasonable to expect that formal learning—that is, schooling—is on the threshold of significant transformation. In

M. Metz · B. Davis (✉)
Werklund School of Education, University of Calgary, Calgary, AB, Canada
e-mail: metzm@ucalgary.ca; brent.davis@ucalgary.ca

S. A. Costa et al. (eds.), *Mathematics (Education) in the Information Age*, Mathematics in Mind, https://doi.org/10.1007/978-3-030-59177-9_12

this chapter, we muse on possible meanings and consequences of that realization, specifically as it pertains to mathematics education.

The social and cultural conditions of this potential transformation are not without precedent. Indeed, there are striking social and technological parallels between current circumstances and the historical moment that saw the original invention of public schooling. At the risk of oversimplification, key motivations for the creation of mass formal education in the western world included a dramatic shift in access to craft and scientific knowledge (enabled by printing presses and postal services), an associated convulsion in knowledge production (enabled by the co-amplifying influences of research institutions and business), an exponential growth of wealth, and the creation of legal systems that gave new rights to the disadvantaged as it recognized the dangers of increasingly inequitable distributions of that new wealth.

The modern school was partly a response to and partly a contributor to these intertwined convulsions. Simultaneously controlling and enabling, mandatory mass education was imposed as much to protect children from an exploitative labor market as it was to equip them with the basic tools needed to contribute effectively to that market. From the start, these basics were identified as the abilities to decode written texts, transcribe simple dictations, and perform uncomplicated calculations—or, more colloquially, "readin', 'ritin', and 'rithmetic." That is, the word "basics" originally signalled *minimal necessary skills* for workers. It pointed to some disciplines, but it said nothing about those disciplines themselves. Unfortunately, as the school became an entrenched and integral aspect of modern culture, the original, context-sensitive meaning of basics was lost. Thus, as society evolved, the basics remained stubbornly resilient. This detail is especially evident in school mathematics where, in the popular arena, "basics" is now typically assumed to refer to adding, subtracting, multiplying, and dividing—that is, not as a set of needs fitted to a particular context at a particular time, but as a reference to an assumed-to-be-natural foundation to mathematics. Indeed, the phrase "learning the basics" is often treated synonymously to "learning simple arithmetic."

Consequently, the construct of basics has become an albatross around the neck of mathematics education. As we develop in this chapter, for example, the notion was as the epicentre of multiple twentieth-century "reform" efforts, which sought to replace the traditionalist emphasis on mastery of *procedural steps* with a focus on *conceptual steps*—that is, to reframe mathematical competence in terms of progress toward deep understanding rather than mastery of technical procedures. That shift was tethered to dramatic developments in psychology and philosophy that contributed to new understandings of learning, which in turn revealed that the beliefs that oriented the original design of public schooling are plainly indefensible. (Even so, they still prevail.)

Profound and consequential insights into learning continue to emerge, now driven principally by the cognitive sciences. In this chapter, we use the notion of *critical discernments* to draw together some provocative emerging ideas and to explore their educational relevance against the now-popular contrast of *procedural steps* and *conceptual steps*. In the process, we also attempt to interrupt the contemporary meaning of "basics" by illustrating our discussions with concepts that we assert are basic to this era, but that are currently given minor attention in most formal curricula. On that matter, we (the authors) are unaware of any mathematics curriculum revision or mathematics teaching reform effort within our lifetimes—anytime, anywhere—that has not been stymied by demands to attend to the "basics." Our hope is that giving heed to matters of basic to whom, when, and where might contribute to efforts toward change.

Companion considerations are the conceptions of "learning" and "teaching" that arose alongside and continue to function in symbiotic relationship with the original notion of basics. Our suspicion, rooted in decades of engaging with classroom teachers, curriculum developers, and policy makers, is that a reason that the tendency to conflate "basics" and "mathematics" is so pervasive and so resistant is that that the assumed relationship is part of a grander flock of associations—that is, of mutually confirming assumptions of the nature of knowledge, the processes of learning, and the mechanics of teaching. In that regard, it appears that efforts to conceive of a mathematics education that is fitted to the moment are complex: A revisionist conception must simultaneously address matters of appropriate content and defensible practice. That is, it must engage with three sets of questions, seeking to understand the conditions of learning (*who*, *when*, *where*, *why?*), to identify and situate content (*what?*), and to define classroom practices based on current knowledge of human cognition (*how?*).

We attempt to take on all three of these matters in this chapter, but in differentiated ways. We start by taking on the *how*—an entrance point that is more intended to uncover some of the intricate web of associations that have over recent decades hobbled intelligent and action-oriented engagements with the other two matters. After that reframing, we turn illustration-based engagements with the conditions and content questions, moving on the conviction that actual experience with a new form of mathematics pedagogy is likely to be more compelling than an academic argument.

Learning: From "Getting" to "Constructing" to "Differentiating"

"Learning" is one of those phenomena that is intimately familiar, but shallowly understood. This point is cogently illustrated through the website, Discourses on Learning in Education (https://learningdiscourses.com), which describes, contrasts, and clusters over nearly 900 (at the time of this writing) perspectives on—that is, metaphors for, definitions of, theories on, strategies of—learning that are represented in the current education literature.

One of those discourses is popularly known as "twenty-first-century learning" or "Deeper Learning"—which, as first hearing, might seem an obvious alignment with the themes of this chapter. A blend of several prominent contemporary discourses, Deeper Learning is explicitly concerned with transforming formal education in ways that fit with emergent personal, social, cultural, technological, and economic conditions. While there are many varieties of the discourse, they tend to cluster around the same set of educational goals (e.g., robust academic outcomes, higher-level thinking skill, positive attitudes, technological proficiency, honed social skills) and to be defined by a specific cluster of teaching strategies (e.g., centered on real-world issues, oriented toward problems that are relevant to learners, choice-rich tasks, access to diverse tools and resources, frequent formative assessments, flexible and frequent opportunities to collaborate).

On the surface, then, Deeper Learning sounds like a movement that is hard to critique. However, even a shallow examination of the discourse reveals that, in fact, it rests on pretty much the same assumptions on knowledge, learning, and education as the traditional, *shallower* approach to schooling that it is presumed to critique.

The learningdiscourses.com site was motivated by this sort of realization. It is designed to assist in making sense of and sorting through competing and complementary perspectives. The project is informed by contemporary research in the cognitive sciences, a transdisciplinary domain that brings together psychology, linguistics, computer science, neuroscience, anthropology, philosophy, and other realms of inquiry. The cognitive sciences focus on the tools and strategies used by humans to make sense of the world, including especially tactics employed to maintain illusions of certainty against a reality of gaping holes in information, frequent flaws in logic, inevitable errors in recall, and implicit prejudices in perception. "Metaphor" figures centrally in these discussions, as both a means and a focus of analysis.

That emphasis is grounded in the twentieth-century realization that human thought is mainly analogical/associative rather than logical/deductive. Much of cognitive science research is thus trained on how metaphoric associations across domains of experience can orient perception, prompt action, bias interpretation, and infuse justifications. That focus turns out to be useful to sort through current discussions in education. As mentioned, the Discourses on Learning in Education site reviews and relates more than 900 currently active perspectives. That number is daunting. Somewhat less daunting, however, is that the number of core metaphors used across these discourses is much smaller (certainly under 50), and fewer than a dozen have any significant traction. As well, major educational movements tend to be associated with specific metaphors.

For example, traditional education is strongly reliant on metaphors through which knowledge is characterized as some sort of stable *object*, by which learning comes to be understood as *getting* that object. The intertwined notions are evident in such phrases as "collections of facts," "gathering of information," "tossing around ideas," "picking things up," "holding a belief," "getting it," "getting to," and "learning stuff." Ancient in origin (see Ong 1982), the grounding knowledge-as-object metaphor can be taken to suggest that there is a real truth, out there, stable, eternal, independent of knowers, untainted, and benign. The cultural priority of these qualities was later amplified in the first scientific revolution, as the ideal of objective truth, and further amplified as a nascent global capitalism found ways to commodify knowledge, creating a marketable thing.

As a means to understand the manner in which humans experience their truths, the knowledge-as-object metaphor has its value. However, as a principle for structuring formal learning, it is lacking. Nevertheless, the cluster of associations that have arisen around this figurative notion has long served as a grounding principle in public education. When knowledge is understood as a set of objects, then it makes sense to conceive of curriculum development in terms of selecting the most-worthy objects and formatting encounters with them. It also makes sense to approach their study as a systematic mastery of their parts. It renders learning a matter of picking things up, packing them in, and bouncing them back. It enables the interpretation of intelligence as how much one can hold—and that highly troublesome notion undergirds a multi-billion-dollar industry focused on measuring these imagined capacities. Ultimately, an uncritical embrace of the knowledge-as-object metaphor defines both learner and teacher, the former as a vessel or recipient, and the latter as a conveyer or deliverer.

The poverty of this cluster of notions was a major focus of psychological research in the early-twentieth century. To bring the issue to the fore, researchers test-drove a variety of new metaphors for learning, with *associating* and *constructing* figuring most prominently. Efforts were made to conceive of learning as an iterative cycle of interpretation, by which one's knowledge was framed as a coherent-but-evolving web of associations. The associated rise of "constructivist" theories among educators in the last half of the twentieth century represented an attempt to format the conversation for educators. Around school mathematics, constructivisms served as the main theoretical engines in major reform efforts, as they were used to alert educators to the problems associated with entrenched-but-invisible object-based metaphors and to the possibilities of taking up action-based metaphors. Problem-solving, personal strategies, learning from errors, talk-aloud protocols, manipulative-based explorations, and a grabbag other constructivism-influenced emphases were soon taken up. One of the popular memes used to collect these new ideas, and to distinguish them with entrenched notions, was a distinction proposed by Skemp (1976) between procedural and conceptual. Procedural was used to tag traditionalist emphases on acquisition and mindless mastery, and conceptual signalled shifts toward construction and making meaning.

Unfortunately, the shift didn't have the impact that theorists hoped. Our suspicion is that the ease with which notions of "constructing" can be blended with notions of "acquiring" proved to be debilitating. That is, the proposed new cloud of associations was perhaps not sufficiently distinct, and so more often than not they were subsumed into established practices and structures. Indeed, even some leading mathematics education researchers seemed to miss the point. Consider Sfard's (1998: 5) conclusion as she contrasted object-based and action-based conceptions of knowledge and learning:

> Concepts are to be understood as basic units of knowledge that can be accumulated, gradually refined, and combined to form ever richer cognitive structures. The picture is not much different when we talk about the learner as a person who constructs meaning.

Further to the surrounding issues, perhaps no one should be surprised that efforts to reform mathematics curriculum have been no more impactful than efforts to reform mathematics teaching. The miscontrual of basics continues in force, a linearized trajectory through prespecified content still dominates.

Enter the cognitive sciences. In recent decades, some educators have started to move to a much more distinct set of metaphors that frame learning in terms of differentiating—a two-layered action of noticing aspects of one's experiences and noticing/construing associations among those noticings. This idea is part of the growing appreciation that humans aren't especially logical. In fact, sapiens are bad at deductive reasoning—and, left to their own devices, tend to fall back on situation-specific and opportunistic tactics to get through situations that would be better managed through systematic thought. Humans extrapolate from past events, they seek patterns in the moment, they impose familiar metaphors, they re-enact established scripts. Thankfully, humans have also learned to off-load the demands of logical thought onto mechanical tools—except, for some reason, in contexts such as most public schools, where there remains an insistence that learners attempt to master mechanical processes that no longer need to be mastered. (To be clear about the point here: We believe that, to learn mathematics, learners must master concepts. But, as we develop, that sort of mastery is quite distinct from the mastery of multi-step procedures.)

What's particularly interesting about the metaphor of learning as differentiating (i.e., noticing and knitting noticings) is its utility for revealing the intellectual poverty of so many educational practices. For example, an immediate consequence of taking this metaphor seriously is that one must be especially attentive to what, exactly, learners are supposed to differentiate, how to channel attentions, how to organize experiences to increase the likelihood of useful associations, and so on. That is, the notion of learning as differentiating takes us immediately to a different model of teaching—one that simultaneously reveals the incoherence of many contemporary educational obsessions while offering a frame for alternative attitudes toward teaching and curriculum. We develop this and associated ideas in the last half of this chapter. But, before getting there, we must take on our second question, on the nature of mathematics. How does the differentiating metaphor prompt us to look at mathematics, and what does mathematical knowledge tell us about how it should be learned?

Mathematics: From "Building" to "Structure" to "Network"

Through the history of modern education, mathematics teaching practice has been consistent with prevailing beliefs about the discipline. For instance, a prominent, and likely dominant, belief is that mathematics is like a building. It has foundations. It has levels, and those levels are ordered. Hence, teaching and curriculum should be attentive to establishing solid foundations and tracing out its levels in logical order. That is, not only is mathematics an object, it is a specific sort of object that dictates topics and orders. Variations on these themes have defined school mathematics since the 1600s, in no small part because they are so compatible with the knowledge-as-object and learning-as-getting metaphors.

The realization that assumptions about the discipline affords an alternative characterization of many efforts to reform school mathematics in the twentieth century. In particular, in the last half of the century, a large number of teachers and researchers who embraced a constructivist sensibility lined up behind a new definition of mathematics—namely, as *what mathematicians do*. Circularity of logic notwithstanding, this shift in definition meshed with constructivist principles of learner agency, inseparability of knower and known, and gradual unfolding of possibility. It also shone a light on appropriate teaching emphases. In that regard, mathematicians were seen to be principally focused on solving problems. Authentic problems. Real problems. Sometimes open-ended problems. This shift in emphasis tied in nicely with progressivist emphases on authenticity and relevance, among other foci. It also fit with emerging sentiments and sensibilities that were later to evolve into Deeper Learning, as described in the previous section.

The move also set up what came to be known as the "Math Wars"—an ongoing, mainly North American-based tension between, in simplest terms, believers in back-to-basics sorts and proponents of problem-solving. Since the late 1980s, the Math Wars have dominated discussions of math teaching practice. For our purposes here, the critical detail is not the explicit tension, but that the Math Wars are enabled and perpetuated by two incompatible sensibilities—that is, two grand flocks of implicit association, each of which is internally consistent, but neither of which is especially defensible. The first of these found its anchor in the assumption that mathematics is an object that exists independent of humanity. In this flock, math learning is about faithfully reconstituting a *fixed reality*. The second flock swirls around the conviction that mathematics is an *evolving structure*—a hallmark of human creativity that emerges when logic is a defining quality. In terms of pragmatic consequences, the structure metaphor grounded criticisms of linear curricula, overly parsed concepts, isolated skills, and procedural steps while it prompted attentions to rich problems, meaningful contexts, flexible sequencing, and conceptual steps.

Yet, somehow, most efforts to enact these notions have not gone well. Somewhat ironically, a likely reason for the failure was anticipated by the person most commonly associated with problem-based learning. Noticing the tendency of humans to frame differences in terms of polarities, more than a century ago John Dewey (1910) cautioned that seeing differences in terms of polar opposites might compel debaters to think that those opposites must bookend all possibilities. That assumption, Dewey (1910: 9) worried, constrained thinking rather than enabling it, as he concluded that, "in fact intellectual progress usually occurs through sheer abandonment of those questions together with both of the alternatives they assume…We don't solve them: we get over them."

For instance, it might be tempting to think that the full spectrum of possibility for school mathematics is captured between "traditional" and "reform" sensibilities. On the one hand, mathematics is seen as pre-determined and pristinely organized—that is, it is regarded as something *discovered*. On the other hand, mathematics is viewed as contingent and subject to human interest and whim—that is, something *created*. Surely the continuum defined by "something *discovered*" and "something *created*" encompasses everything.

In fact, however, almost everyone who has framed their thinking with the dyad of "something *discovered*" and "something *created*" has missed a blindingly obvious detail: both elements of the dyad assume a *something*. Both are indexed to an assumption that mathematics is a thing—and, not somewhat ironically, this detail shows up most powerfully around the notion of *discovery*. As mentioned, among Traditionalists, mathematics is usually seen as out there, in the real word, and therefore discovered. Once discovered, however, it makes sense to convey it. Among Reformists, mathematics is most often assumed to be created, but somehow that conviction leads to strong recommendations for discovery-oriented teaching—revealing that object-based assumptions on mathematics have not been jettisoned at all. Perhaps that is why, even though Reformists managed to awaken educators to learner agency by attending to what mathematicians do, the contents and outcomes of most mathematics curricula are scarcely discernible from pre-reform versions, even after a half-century of Reformist influence.

Getting over the Math Wars, then, may be dependent on a compelling and defensible alternative to the implicit-but-pervasive knowledge-as-object metaphor that continues to undergird almost all discussions of school mathematics. Fortuitously, many alternatives have been developed, especially over the past few decades. One that we find especially useful is the metaphor that k*nowledge is systemic coherence across levels of organization*, from which it follows that learning might be associated with making and acting on differentiations that enable systemic coherence. That is, learning is about noticing and knitting noticings—and, in turn, that blend sets up a model of school mathematics that aligns with neither side of the Math Wars. And it doesn't land between them either.

Over the past few decades, in efforts to understand the nature of their discipline, many mathematics researchers have turned its tools onto the discipline itself. A consistent conclusion is that mathematics is a complex system (e.g., Foote 2007). Consequently, mathematics has a decentralized network structure. Other phenomena that have this structure include cultures, ecosystems, and brains. More pointedly, mathematics is *not* an object—and, with that, the premises of the Math Wars crumble. As do both Traditionalist and Reformist teaching.

So, how might an educator approach mathematics when knowledge is understood as systemic coherence across levels of organization? To answer that question, we focus not on the elements of mathematics but on how the elements of mathematics might be made available in learners' experiences. A decentralized network comprises both nodes and links—which, in the case of mathematics, have been associated with "principles" (i.e., stable aspects of existence, such as patterns and forms) and "logics" (i.e., different means of combining principles into systems of interpretation). Learning mathematics, then, is about differentiating—that is, becoming aware of principles (i.e., noticing) and applying logics (i.e., knitting noticings).

Correspondingly, teaching comes to be about channeling attentions and juxtaposing experiences to support appropriate linking.

And that takes us to a model of mathematics teaching informed by "Variation Theory," which we argue is fitted to this Information Age.

Why Variation Theory?

As the Math Wars have continued to polarize discussions regarding the best ways to teach math (Schoenfeld 2004), it is clear that we have not yet adequately answered Chazan and Ball's (1999) call to go "beyond being told not to tell." Marton's theory of variation (Marton and Booth 1997; Marton and Tsui 2004; Marton 2015) offers powerful insights that allow a clear alternative to both telling and discovering—and to the knowledge-as-object metaphors upon which they are based. While it is impossible to "transmit" understanding or "process" the products of perception, it *is* possible to offer deliberate contrasts that dramatically increase the likelihood that learners will perceive intended principles and relationships in a particular way. In addition to effective prompting techniques, this requires careful attention to both short and long-term structuring—which we describe with the contronym, *raveling*—of mathematical ideas to which we might prompt. Neither effective prompting nor long-term raveling feature prominently in discussions of traditional vs. reform approaches, which has likely contributed to the longevity of the Math Wars. Traditional methods *can* work. So can reform methods. But the alleged reasons they work (or reasons the other does not) may have more to do with elements of pedagogy that do not even enter the conversation; further complicating matters, success may be *mis*attributed to one or other "contemporary obsession" (Preciado-Babb et al. 2020).

In this part, we further develop the key ideas underlying variation theory and relate them to principles of variation pedagogy developed independently in China (Gu et al. 2004; Huang and Leung 2004; Lai and Murray 2014). Following that, below we offer an interpretation of variation theory that integrates Marton's theoretical principles and Chinese pedagogical principles into a nested set of variation *types* that we've found helpful for designing short and long-term pedagogical sequences that support fractal awareness consistent with the nature of mathematics.

Learning as Differentiation

Marton's theory is based on the premise of *learning as differentiation*, which might be contrasted with *learning as enrichment*; i.e., that perception is always necessarily partial, depending on what we separate from an undifferentiated whole. Emerging from this is a distinctive principle that lies at the heart of Marton's work: The *Principle of Difference* states that we discern new ideas when they are contrasted against a constant background. *Highlighting difference to prompt distinction-making* is itself clarified by contrasting it with the common practice of (attempting to) *highlight similarity to promote association-making*. While association-making is indeed important to learning, Marton argued that we do not discern *new* meanings by perceiving similarities among examples that otherwise differ—i.e., through induction. If we can't perceive something in one place, we won't see it in two, or three, or a hundred. We can, however, *generalize* similarity among *previously* discerned examples (we also make metaphorical associations, but we will take that up a bit later).

Distinguishing generalizing from induction is essential to understanding and effectively using variation. This can be tricky, because often the patterns of variation used to prompt generalization are the same ones we might be tempted to offer in the hope of prompting induction. But order matters: Once separated through contrast, ideas become perceptible and can *then* be generalized. The examples may be the same, but the manner in which associations are assumed to be made is not. What can be highlighted via contrast is also constrained by prior knowledge, but in a different manner: We can't simply prompt attention to advanced mathematical ideas unless the necessary contrasts have themselves been sufficiently prompted. Thus, mathematical ideas must be carefully "raveled" so that we prompt attention at a level where learners are able to make sense of offered contrasts. It turns out that the levels can themselves be usefully described in more general terms, which is an important elaboration of the variation types that we discuss below.

The Principle of Difference: Induction vs. Generalization

The Principle of Difference is both less intuitive and more powerful than often appears at first glance. When we are trying to explain something that is familiar to us, it can *seem* as though multiple examples *should* support deeper insight. This is likely because once something has been discerned, multiple examples *can* add clarity through expansion of the example space associated with that idea (Watson and Mason 2005, 2006). Again, however, this is about *generalizing* existing understanding. If the particular something hasn't yet been discerned, it's impossible to simply *induce* what the many examples are examples *of*: We can't *induce* meaning from similarity if a unifying feature is unavailable.

Mason, Burton, and Stacey (1982/2010) recommended "generalizing and specializing" as a way of seeking deeper insight; it is through exploring variations of a particular case that we often find insight into both *that* case and *a more general class of cases* to which it belongs. Importantly, however, this is not about *finding similarities among varied examples*, but about *finding the perturbations under which the broader category remains intact* (which may itself be influenced by the particular conditions of investigation).

In some cases, a single contrast can provide the necessary insight for generalizing a particular feature, which is why sometimes it's possible to "see the general in the specific" (Mason and Pimm 1984; Watson and Mason 2006): Doing so involves seeing certain parameters of a problem as *variable*, which provides the necessary contrast for generalizing. In summary: *separation* (through *contrast*, not induction) must precede *generalization*; further, the very same examples that are inadequate for separation are ideal for generalization. Once separated and generalized, different features may be simultaneously varied, or *fused*.

Separating, Generalizing, and Fusing

What exactly does it mean to separate through contrast? If we want to highlight the meaning of, for example, *apple*, highlighting difference would involve *contrasting* an apple with other things that are like apples in as many ways as possible but differ with respect to some essential feature. At the moment of discerning, both *apple* and the broader whole from which it has been separated (Food? Fruit? Spherical objects?) become namable, but these namable "things" are less important than the un-namable *difference* that defines them. In other words, the notion of difference is essential to transcending the knowledge-as-object metaphor that has proved so intransigent over decades of attempts to improve math education.

Once the notion of apple has been separated from a background—say of food, or fruit, or spherical objects (i.e., once *apple* becomes a discernible and thus name-able difference)—we may *then* generalize to a broader class of apples. Although this class may be *defined* in terms of what all apples have in common, it is generated and bounded through expansion of allowable differences (Can it still be an apple if it has pink flesh?). To generalize, we hold the notion of apple constant and consider the permissibility of particular variations of apple—which looks just like the pattern of induction mentioned above, except that we're using *difference* (not similarity) to test the limits of our definition of apple. In other words, generalizing is about perceiving *differences between differences* (i.e. variations of *apple*, which is itself distinct from *non-apple*)—and thus could be considered the sort of level change that lies at the heart of the hierarchy we are proposing.

According to Marton (2015), new ideas must be prompted in a manner such that what is general and what is specific are discerned *simultaneously*. When distinguishing apple from non-apple, it may be that a learner becomes aware of the category *fruit* of which apple is a particular type; Marton would call such a category a *critical aspect*. If apple is the first fruit to be so separated from the broader category, both *fruit* (as a category) and *apple* (as one member of that category) are perceived at the same time. In this case, the contrast highlights both the apple and a hierarchical structure involving apples and fruit (i.e., both the critical *aspect* fruit and the critical *feature* apple). Thus, "What is general and what is specific are discerned simultaneously when a new meaning is appropriated. There cannot be any features experienced without the awareness of the aspect that unites them, nor can there be any aspect experienced without the awareness of features that belong to it. Differences and features that differ cannot exist without each other" (Marton 2015:

48). Once apple and fruit have been thus separated, we might be moved to lay out the particular criteria we see as essential to each. Having done so, we might hold up *new* examples to those criteria and thereby classify them as fruit, apple, or both. *Classification* is distinct from *generalizing* in that it sets particular features of a particular example against an articulated definition; generalizing, on the other hand, takes particular features of a particular example and identifies a space of possibility bounded by the constraints of the experiential context as opposed to by the defined criteria of a definition.

Watson and Mason (2005) helpfully referred to critical aspects in terms of *dimensions of possible variation* and critical *features* in terms of *range of permissible change.* To highlight allowable variations of apple, apple becomes the critical aspect, which can be generalized according to variation in (familiar) features such as color, shape, size, and flavor. In doing so, it is helpful to contrast and vary features one at a time: Apples can be various shades of yellow, green, or red. They range from quite round (Macintosh) to a bit lumpy (Red Delicious). They can be smaller than a tennis ball (crab-apple) or as big as a softball (Fuji). They can range from quite sour (Granny Smith) to very sweet (Fuji). Color, shape, size, and flavor are *critical features* (of the *critical aspect* apple), and each can be varied within certain parameters. Similarly, "yellow, green, and red" are features of the aspect color, but discerning color itself isn't the focus at this time; in the context of the apple, color is assumed as prior knowledge.

Even when critical features have been carefully *named* in an attempt to separate them for attention, teachers and resources frequently attempt to prompt attention to the *named thing* rather than to relevant *differences*—differences which might *then* be given a name. This is how easily Marton's induction insinuates itself into pedagogies where the metaphor of knowledge-as-object has not been interrupted.

To offer a simple example, "practice rounding" tends to involve practice sets that cluster varied numbers to be rounded to nearest ten, then numbers to be rounded to the nearest hundred, and so on. Direct contrast between rounding to the nearest ten and nearest hundred is thus not easily perceived. Alternatively, the pattern of variation in Fig. 12.1 offers *direct juxtaposition* of rounding to the nearest ten, hundred, thousand, and ten-thousand—and requires learners to round *the same number* to varied place values (note that it's important that students are prompted to work from top to bottom, as moving left to right offers the same pattern of variation we're hoping to interrupt). These contrasts offer meaningful information about the impact of rounding. *More* carefully chosen variation of *fewer* examples can be very powerful, because learners must practice *making key discernments* rather than merely practice

Round to nearest:	3421	5421	13421	16476
10				
100				
1000				
10000				

Round to nearest:	1329	1389	1789	6789
10				
100				
1000				
10000				

Round to nearest:	1119	1199	1999	9999
10				
100				
1000				
10000				

Fig. 12.1 Rounding

a procedure. The contrasts *between* selected numbers and between successive charts are also carefully chosen, but these differences between differences can only make sense if the first order of difference has been successfully prompted. If we treat the first level of difference as a *thing* rather than a difference, every level thereafter becomes inaccessible. In this sense, first-order differences become the essential criterion for what is truly *basic* to structurally coherent mathematics pedagogy.

Similarly, when learning to add multi-digit numbers, learners may be asked first to add without regrouping, then to add with regrouping. A key discernment in doing so, however, is recognition that tens in the ones' place *can* be re-grouped. In other words, re-grouping can only be perceived in contrast to *not* re-grouping (and vice versa). When examples and practice sets separate adding with and without re-grouping, such contrast is far less obvious. Alternatively, if *varied tens* in the ones' place are directly juxtaposed (and include *no* tens), particularly in a manner that highlights those tens, the meaning of re-grouping is more likely to be perceived see Fig. 12.2). At the very least, learners should have to decide whether or not to re-group (zero tens or one ten). In practice sets where every example involves trading a single ten, we have observed learners go through the entire set and placing a "1" in the box for trades. They are not *distinguishing* between trades and no trades (or between one ten and other numbers of tens); they are merely performing a step that is highly limited to the particular context of that practice set. Note that in the last example, the lack of color-coding, slight mixing of the ten pairs, and inclusion of an extra one introduce elements that widen the space in which learners are expected to make appropriate discernments. Depending on learners' background and confidence, these features might be introduced one at a time.

To summarize: Once an idea is *separated* through contrast, essential features may also be highlighted by varying them against a constant background and *generalized* by identifying *non-essential* features. They may also be *fused* by co-varying previously discerned (and possibly generalized) features. Returning to the apple,

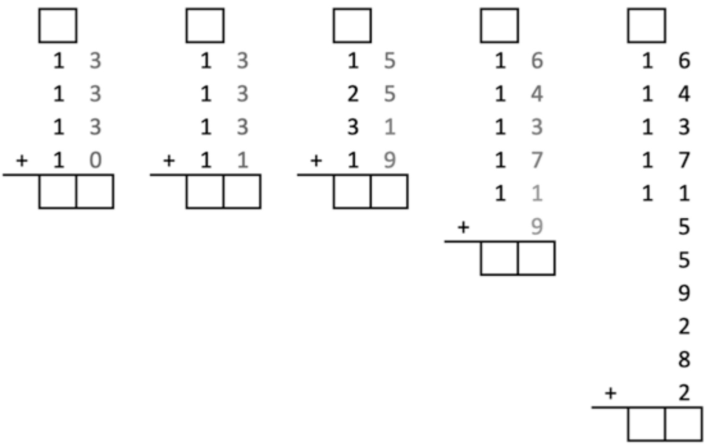

Fig. 12.2 Re-grouping tens

eventually, we recognize *apple* even when multiple features co-vary and even when those features co-vary in ways that produce novel situations. Even if we were suddenly confronted with a crab-apple-sized, Red Delicious-shaped, yellow-coloured apple, we would likely recognize it as an apple. Similarly, we can learn to add multi-digit numbers that involve re-grouping into *any* place value and with *any* number remaining in each place value. More significantly than what we can *do*, however, is that we may now *recognize* the very idea of re-grouping tens in a more broadly generalized sense that makes it available in a wide variety of other contexts, including all of the other so-called "basic" operations. Note that it's not the operations themselves that are basic, but the critical discernments upon which the traditional-defined basics need to be based.

While separating, generalizing, and fusing have to do with discerning the effects of multiple co-varying features, they do not fully account for how we generate mathematics or how we learn mathematics: Mathematical knowledge also expands both through *abduction* and through the *integration* of different ideas.

Abduction

If mathematical knowing has to do with organizing information into accurate hierarchies and identifying the *logical* implications implied by those structures, it may seem that mathematics is inherently a logical endeavour. But the *formation* of those hierarchies is an abductive—or more specifically, an analogical—process, which is not surprising considering that such hierarchies are created by analogical minds. Here, the similarity we decried in the context of induction assumes a prominent role, though difference is still essential to prompting new meaning. We generalize when we compare examples and decide whether identified differences are consistent with pre-specified criteria; we abduct when we transfer explanatory structure (consciously or unconsciously) between that which we perceive as similar. We *analogize* when we consider the appropriateness of transfer and do so (or refrain from doing so) intentionally.

To continue with our fruit example, it may be that a learner has already discerned apples and oranges but not considered their relationship. Doing so may simply involve combining them into a single category based on *specified* and *familiar* criteria; as noted about, this is *classification*. However, it may involve consideration of whether recognized shared features allow transfer of explanatory structure from one to the other. If so, we are talking not just about classification, but about *abduction*. Although difference is required to separate *new* ideas, the human mind is adept at perceiving similarities among previously discerned ideas. Abduction is *not* about seeing something *new*—it's about recognizing something familiar in a different context. From there, it is possible to consider whether what is known about each context may usefully or accurately be brought to bear on the other, though this aspect of abducting appears much less intuitive. Humans are deeply prone to jumping to conclusions based on perceived similarities. In any case, difference can only be perceived between ideas that have been previously discerned; i.e., differences can only be perceived between previously perceived differences. This would imply infinite regress if we assumed a blank-state infant mind, but we know that humans come into the world already attuned to particular differences from which all others eventually evolve (Lakoff and Núñez 2000).

Separating, generalizing, and fusing contribute to effective knowledge hierarchies. Marton (2015) addressed the importance of such hierarchies (particularly in the context of writing). Watson and Mason (2006) further emphasized the fractal nature of those hierarchies and highlighted the role of abstraction—which is consistent with what we're calling abduction—in their formation:

> However, to make mathematical progress the results of the images, models, and generalizations thus created have to become tools for more sophisticated mathematics. We see generalization as sensing the possible variation in a relationship, and abstraction as shifting from seeing relationships as specific to the situation, to seeing them as potential properties of similar situations. (Watson and Mason 2006: 94)

Taken together, the rejection of induction and the articulation of the role of difference in generalizing and abducting further support the importance of abandoning *mental things*. Doing so also helps resolve an apparent paradox highlighted by Watson (2017): If much of mathematics is defined in terms of *similarities* (defined by dependency relationships among variables) how can we use difference to prompt to similarity? One way of looking at this is that similarity is always between *similar differences*; if not, we couldn't have perceived them in the first place.

Integration

When used with well-raveled content, variation *theory* can offer powerful insights that contribute to effective variation *pedagogy*. To do so effectively, we must of course be clear about what we want to prompt attention to. But this is not as straightforward or intuitive as it might seem, particularly in the context of mathematics education. Marton (2015: 176) emphasized that "the object of learning that is used as a lens for inspecting the teaching may or may not be identical with the intended object of learning (i.e., the learning that the teacher had hoped to contribute to)." In mathematics it is frequently the case that instrumental learning is mistaken for relational learning (Skemp 1976). Carefully generated patterns of variation intended to teach particular mathematical ideas will ultimately fail if those ideas are defined only in instrumental terms; i.e., as *things* rather than *differences*.

Many who focus on step-based approaches emphasize that they *do* focus on the meaning of those steps. Even when the focus is on the underlying conceptual meaning of those steps, however, a procedure does not always offer an effective means of raveling the mathematical ideas required to make sense of that procedure. Many procedures require the integration of multiple ideas that, if not previously discerned and generalized, are very difficult to integrate (see Fig. 12.3).

Steps vs. Discernments

How does the standard algorithm for long division work?

Procedural Steps

1. Figure out how many times the divisor fits into the digit with the highest place value in the divisor. Write the number of times the divisor fits above the corresponding digit in the dividend (ignore remainder). This is the first digit in the quotient.
2. Multiply the first digit in the quotient by the divisor. Write the product beneath the first digit of the dividend.
3. Subtract.
4. Bring down the next digit of the dividend, and write it beside the number you got when subtracting. This will form a two-digit number.
5. Divide the two-digit number by the divisor, and write the answer you get beside the one you got in Step 1.
6. Multiply the second digit of the dividend by the divisor. Write the product beneath the two-digit number you divided in Step 4. Subtract.
7. Repeat Steps 2-4 until all digits have been divided (Divide, Multiply, Subtract, Bring Down, Repeat)

Conceptual Steps

1. Divide each place value one at a time. Starting with the highest place value, divide it into the number of groups specified by the divisor. Write the quotient above the digit you're dividing.
2. Multiply that answer by the divisor to find out how many of that place value have now been placed. Write the total beneath the corresponding place value in the dividend.
3. Subtract the number you got in #2 from the digit in the dividend that you were working with. The difference is the remainder that's left to be divided.
4. Combine the difference in #3 with the digit in the next place value; you can do this by simply bringing the next digit down. Unless Step 1 had a remainder of zero, you will now have a two-digit number.
5. Divide the two-digit number you found at the end of Step 3 by the divisor. Write the quotient beside your answer in #1.
6. Repeat Steps 2-4 until you have divided each place value. If there is still a remainder when you're finished, you can just state how many are left over.

Fig. 12.3 Procedural steps vs. conceptual steps (long division)

Steps vs. Discernments

How does the standard algorithm for long division work?

Conceptual Steps

1. Divide each place value one at a time. Starting with the highest place value, divide it into the number of groups specified by the divisor. Write the quotient above the digit you're dividing.
2. Multiply that answer by the divisor to find out how many of that place value have now been placed. Write the total beneath the corresponding place value in the dividend.
3. Subtract the number you got in #2 from the digit in the dividend that you were working with. The difference is the remainder that's left to be divided.
4. Combine the difference in #3 with the digit in the next place value; you can do this by simply bringing the next digit down. Unless Step 1 had a remainder of zero, you will now have a two-digit number.
5. Divide the two-digit number you found at the end of Step 3 by the divisor. Write the quotient beside your answer in #1.
6. Repeat Steps 2-4 until you have divided each place value. If there is still a remainder when you're finished, you can just state how many are left over.

Critical Discernments

CD#1: Division can be thought of in partitive or quotative terms. Here, we will focus on partitive division.

CD#2: When dividing, you can break the dividend into parts, divide each part into the number of groups specified by the divisor, then combine them (distributive property).

CD#3: If you break the dividend into parts that divide evenly by the divisor, only the final group (if any) will have a remainder. If you don't, the remainders from each group will also need to be combined. The combined remainders may be large enough to further divide.

CD#4: The standard algorithm for long division uses a special case of this strategy whereby the parts being divided are specified by the digits in each place value.

Fig. 12.4 Conceptual steps vs. critical discernments (division)

One of the obstacles that some seem to have with understanding critical discernments is that they think so long as steps are explained conceptually, they qualify as critical discernments. This ignores the importance of raveling: Often explaining a step in an algorithm involves multiple discernments (see Fig. 12.4), which is why many learners don't follow the conceptual explanation and beg to simply be given a list of steps. Critical discernments are raveled over time so that learners have made the necessary discernments that allow them to make sense of each new discernment.

Sometimes raveling algorithms involves integrating seemingly unrelated understandings that come together due to their role in the solution to a problem, but a well-raveled sequence should also have elements of progressive differentiation. In math, this often shows up in the form of seeing something as a special case of a broader principle. For example, the long division algorithm is a special case of separating-to-divide (CD #2 in Fig. 12.4), which itself is a refinement of the distributive property. Contrasting different ways of separating a number to divide opens "ways of separating" as a dimension of possible variation while simultaneously expanding the example space consisting of those ways.

Once again, generalizing and specializing emerge as two sides of the same coin. We generalize by varying and identifying boundaries for variation, not by looking for similarity among multiple examples. Nonetheless, it remains important to remain attentive to *implications* of such insights on both (or multiple) levels of the hierarchy and to how they might be elaborated; i.e., to what can be articulated in general terms and to the specific cases that comprise, limit, or extend articulated generalizations (Mason et al. 1982/2010).

This manner of viewing generalizing and specializing has implications for the current obsession with multiple strategies (Preciado-Babb et al. 2020). Rather than learning a variety of different ways to divide and then considering how they're *alike* (Marton's induction), we can progressively refine critical discernments about the nature of multiplication, division, and the distributive property. Each of the typical methods for division emerge from these critical discernments and are thus *already connected*. Again: *The general is recognized at the same moment that the particular is differentiated*; we see the general in the particular (Mason and Pimm 1984) *precisely when the particular emerges through differentiation*. Or, as Marton (2015: 48) put it: "There cannot be any new features experienced without the awareness of the aspect that unites them, nor can there be any aspect experienced without the awareness of features that belong to it. Differences and features that differ cannot exist without each other."

The distinction between procedural and conceptual steps is perhaps even more striking in the case of relating prime factors and factors (see Fig. 12.5). Offering a conceptual explanation of the procedural steps listed here would be grossly insufficient for most learners, because they would require a *sub-ravel* (and a sub-ravel of the sub-ravel) for each step before such an explanation could make sense.

Integrating multiple ideas is distinct from both fusion and from discerning dependency relationships among variables. It has to do with bringing diverse mathematical ideas to bear on a single context or problem, as in modeling and problem solving. Elsewhere, we have used a braiding metaphor (Preciado-Babb et al. 2020) to describe this difficulty: Learners need to braid the strands (i.e., each of the critical discernments on the right) before they can effectively attend to the rope itself. In other words, they would have to braid the strands at the same time that they're braiding the rope. This is also true of the discernments pertaining to long division, but the layers of sub-ravel requiring attention for the unknown to become perceptible may seem less daunting.

Steps vs. Discernments

How do prime factors determine number of factors?

Procedural Steps

1. Find the prime factors of the number.
2. Write the prime factors with exponential notation.
3. Add 1 to each exponent from Step 2. Multiply those numbers together to find the number of factors.

Critical Discernments

CD#1: Every number can be written as the product of prime factors.

CD#2: Every number has a *single* prime factorization.

CD#3: Every number has a *unique* prime factorization.

CD#4: Prime factors combine to make factors.

CD#5: The number of factors a number has is determined by the number of ways you can combine its prime factors.

CD#6: If a number has one prime factor, repeated multiple times, the combinations are varying numbers of that factor.

CD#7: If a number has two or more prime factors, clusters of one prime multiplied by clusters of the other(s) create *new* factors of that number.

CD#8: The *number* of factors can be found by multiplying the number of possibilities for each prime factor in a number's prime factorization.

Fig. 12.5 How do prime factors determine the number of factors?

Conceptual and Procedural Variation

While there is value in talking about the theory of variation more generally, it is illuminating to consider how its principles are specifically implicated in the teaching of mathematics. In China, variation *pedagogy* particular to mathematics (Gu et al. 2004, 2017; Huang and Leung 2004; Lai and Murray 2014) was originally developed independently of Marton's theory, though collaboration among researchers from the two traditions has become common as researchers have recognized their complementarities. Nonetheless, there are some notable differences in emphasis (Huang et al. 2016; Pang et al. 2016; Watson 2017). In particular, Chinese variation pedagogy includes an explicit focus on *sameness* as well as *difference* (Gu et al. 2004). Watson (2017: 85) emphasized that much of mathematics takes *invariant* "dependency relationships" as its object of learning; in other words, similarity is, in fact, essential to mathematics. But earlier we insisted that humans are attuned to difference, not similarity, and that induction does not work. What's going on? We have found that the manner in which *sameness* is prompted is consistent with Marton's principle of difference, but it's important to consider more closely what it means to *use difference to prompt attention to similarity*—or more specifically, to the *underlying dependency relations that generate that similarity*. Prompting to relationship is much different than prompting to pattern (Hewitt 1992).

Gu et al. (2004) distinguished two types of variation important to teaching mathematics: (a) conceptual variation (CV) and (b) procedural variation (PV). Here, "conceptual" and "procedural" are used differently than is typical in Western contexts, where conceptual is roughly synonymous with Skemp's (1976) "relational," and "procedural" is roughly synonymous with his "instrumental." Both conceptual and procedural variation are about sense-making, and they are neither opposed nor competing. According to Gu, Huang, and Marton, conceptual variation offers examples, non-examples/counter-examples, and non-standard examples of a concept; thus, conceptual variation is roughly akin to Marton's contrast and generalization and to Watson and Mason's (2005, 2006) "example spaces." Conceptual variation offers the initial differences that bound a context for learning; it is concerned with clarifying and broadening the space of variation encompassed by a particular idea. Hewitt's (1999, 2001a, b) consideration of what is arbitrary and what is necessary in mathematics further highlights the importance of distinguishing what can and cannot be relationally defined and of teaching in ways that are consistent with this distinction. Even what is arbitrary, however, must be separated by prompting to *difference*; a particular definition or premise may be arbitrarily chosen, but once chosen, it cannot be arbitrarily prompted.

Procedural variation is further differentiated into three sub-categories: (a) varying the features of a problem (PV1); (b) comparing methods for solving a problem (PV2); and (c) considering how a single method can be applied to similar problems (PV3).

The procedural variations are collectively described as "progressively unfolding mathematics activities" (Gu et al. 2004: 319): "[P]rocedural variation intends to pave the way to help students establish the substantive connections between the new object of learning and the previous knowledge" (Gu et al. 2004: 340–341). In this way, successive examples may be experienced as "easier," but this is a particular kind of easier: They make it easier to make significant mathematical discernments, not just easier to complete a practice set or do a certain type of question. While some might see repetitive practice in a set of tasks designed with procedural variation in mind, it in fact involves deliberate change against a constant background; i.e., it's not the repetitive practice of a procedure but a way of highlighting relationships between particular mathematical variables. Lai and Murray (2014) argued convincingly that failure to distinguish between these two types of repetition likely lies at the heart of the perceived "Chinese paradox" (Huang and Leung 2004), whereby Western observers have sometimes struggled to make sense of how allegedly weak Chinese pedagogy consistently produces such strong results.

Differences based on logical hierarchy are also an important consideration when considering effective patterns of variation. A lesson (or a text) is experienced chronologically, but for learning to be effective, ideas offered within that chronology must be structured hierarchically; doing so involves prompting awareness of particular hierarchical relationships and how they are woven into increasingly dense and elaborated webs of association. Marton (2015) reported on several studies (outside of mathematics) where each new awareness was connected to a broader context. In such cases, learning was more effective than in cases where teaching was structured in linear sequence.

The different forms of variation can be seen in terms of a natural hierarchy with the potential to support the sort of meaningful long-term structuring of mathematics envisioned by Dienes (1960). Typically, however, *descriptions* of variation pedagogy often involve unrelated examples used to exemplify types of variation (Kullberg et al. 2017; Sun 2011; Wong 2008). While such studies are useful for prompting attention to the significance of fine-grained variation, it is not easy (a) to see the distinguishing features of different types of variation or (b) to see how they might be used to progressively elaborate an idea beyond the immediate context of a question set or a lesson. In China, coherent long-term raveling may be more clearly supported by carefully developed teaching resources, but elsewhere this is not always the case (Jianhua 2004; Bajaj 2013). Ma (1999) found that Chinese teachers, even those with less formal education than their counterparts elsewhere, showed more profound understanding of the elementary math they were teaching. In such cases, it is particularly vital that progressions highlighting hierarchical structure be embedded in quality resources that support both teachers and students in discerning complex webs of relationships.

Summary

We opened this part by noting that mathematics education has yet to articulate and put into broad practice a meaningful response to Chazan and Ball's (1999) observation that we need to offer teachers more than a directive not to "tell." Variation theory offers a way out of the apparent contradictions that emerge from many of the traditional vs. reform debates. It is not without potential pitfalls, however. Particular contradictions emerge when we attempt to use variation theory in conjunction with the knowledge-as-object. In such cases, variation theory tends to be distorted in one of two ways, depending whether the distortion occurs in a Traditionalist context or in a Reformist context.

Traditionalists may take variation as a means of offering gentler progressions and minimizing cognitive load. In other words, the subtle changes between questions are seen primarily in terms of gentle steps rather than in terms of meaningful contrasts deliberately chosen to make particular differences visible. While it is indeed important to attend to the limits of working memory, effective variation is about increasing clarity, not about making things easier. In fact, attempts to simplify by focusing on one thing at a time often result in the loss of the very contrasts necessary for effective prompting through variation.

Reformists working with the hope that learners will independently discover relevant knowledge-objects often fail to consider the hierarchy of differences required to make necessary contrasts perceptible in the first place. Open problems *can* offer spaces in which the variation of critical features assumes relevance (Runesson 2005), and learners can indeed take responsibility for generating their own patterns of variation emerging from those features (Watson and Mason 2005)—*if* they've discerned relevant dimensions of possible variation, *if* they have appropriate mathematical tools (developed through their own ravel) to manage that variation, and *if* they are not expected to weave the strands they are braiding, so to speak. Not all "paths" in the Reformists' journey follow a sequence that effectively supports attention to necessary differences, the fusion of multiple variables or the integration of prior knowledge.

When variation theory is paired with the metaphor of knowledge as a decentralized network, however, the importance of separating relevant features from an integrated whole and then re-integrating them in a manner that highlights an appropriate web of associations is much easier to talk about; i.e., the metaphor invokes both relationships and language that productively enable thinking and communicating about learning. Here, the significance of the particular ways in which mathematical ideas are raveled assumes prominence: Careful contrasts and wide spaces of variation typically open rich spaces of conceptual variation that subsequent procedural variations may continue to elaborate.

We have found it somewhat challenging to highlight how such hierarchies unfold in a longer-term ravel. To focus on the fine-grained distinctions significant at a particular level makes it harder to step back and focus on the relationships between levels in the hierarchy. To do so, we do not offer the same level of detail *within* each level that we might otherwise do, though we hope that the particular examples we've chosen sufficiently highlight the importance of fine-grained distinctions. Once again, the need to simultaneously attend to an intricate web of understanding at both the immediate and the long-term level fuels our insistence that a carefully developed resource is essential to the coherent, long-term elaboration of mathematical ideas.

Mathematics as Levels of Variation

We have found it helpful to conceptualize *types* of variation in terms of *levels* of variation defined by varied interactions among successive levels of difference (see Fig. 12.6). Level 1 separates and bounds key ideas with which we wish to work. This typically invokes what Hewitt (1999, 2001a, b) deemed the *arbitrary*. Levels 2–5 involve qualitatively different interactions among identified parts, each of which involves *necessary* implications (as opposed to arbitrary definitions) of the ideas established at Level 1. Levels 1–4 form a hierarchy of types: Level 1 uses

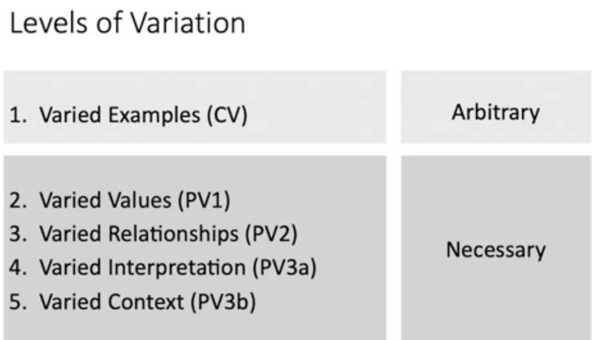

Fig. 12.6 Levels of variation

contrast to separate and generalize key features, Level 2 explores co-variation of those features, Level 3 contrasts relationships among different ways of co-varying (sometimes in the form of sequences and strategies), and Level 4 contrasts relationships among relationships. Level 5 focuses on interactions among previously identified features and relationships, including those that go beyond the object of learning identified in Levels 1–4.

Consistent with our claim that seeing knowledge as a decentralized network significantly influences how we make sense of teaching and learning, these levels might be helpfully compared with Bateson's articulation of logical types pertaining to the development of living things:

1. The parts of any member of Creatura [i.e., living things] are to be compared with other parts of the same individual to give first-order connections.
2. Crabs are to be compared with lobsters or men with horses to find similar relations between parts (i.e., to give second-order connections).
3. The *comparison* between crabs and lobsters is to be compared with the comparison *between* man and horse to provide third-order connections (Bateson 1979/2002: 10).

If we substitute ideas for organisms, we come very close to the framework we are attempting to describe. Hence we move from what Bateson referred to as *serial homology* (where each part within a particular organism is constrained during embryonic development by the previous parts) and *phylogenetic homology* (where new developments are constrained by shared evolutionary history) to what might be considered *parallel homology,* where parts do not act directly upon one another but may yet act in similar ways due to a history of evolving to meet similar evolutionary constraints. The three points above correspond to our Levels 2–4. Within and between each level, information is defined by *difference,* which is precisely why it can't be a *thing*. Difference exists in the space between. Each of the levels (1–5) are however, bounded—by what Bateson referred to as *context* and described in terms of a story, or pattern through time, that links varying elements in a space of shared relevance: "Any A is relevant to any B if both A and B are parts of components of the same 'story'" (p. 14).

Change and Choice

Before we articulate the levels themselves, it's important to take a closer look at the importance of change and choice. Teaching involves prompting attention to relevant differences, which are themselves essential for prompting attention to associations. Prompting has two key elements: what we offer (change) and what we invite learners to distinguish (choice). We must *offer* relevant contrasts and *invite* particular noticings, and to do so, we must offer tasks that require learners to *make relevant distinctions*. Attention to "change and choice" in this manner has a profound impact on what we call "practice:" Most of what gets called practice is "practice doing," but *good* practice is "practice discerning." In addition, examples and tasks must each use *appropriate* and *sufficient* contrast; i.e. they should be sufficiently different, uncluttered by distractors, and clearly juxtaposed in both time and space.

Finally, it is essential to acknowledge learners' roles in making themselves receptive to change: It is by moving our heads that our eyes receive differential signals regarding the position of objects in space and that we may thus perceive depth; this is also how our ears receive differential signals regarding the direction of sound, that we may thus perceive the direction from which a particular sound emanated; and it is by dragging our fingers over a surface that we may perceive differences that alert us to the shape or texture of whatever we are touching. Teachers may prompt to the significance of difference by inviting attention to relevant contrasts, but as learners come to *expect* such differences in the sequence of examples and tasks they are offered, they learn to do the mental equivalent of moving their head or dragging their finger, now over a *set of ideas*, but still with the aim of detecting the *differences* that contain relevant information. In the remainder of this section, we illustrate the five levels in a manner that we hope offers an abbreviated experience of relevant differences. A more elaborate sequence with many more opportunities for engagement may be found in Unit 3 of the Math Minds Online Course (Math Minds 2020).

Level 1: Bounding and Naming Differences

Level 1 defines the nature and limits of what we want to work with; i.e., to the particular aspects and features with which we wish to work. Level 1 is akin to conceptual variation in Gu et al.'s (2004) description of Chinese pedagogy and to separating in Marton's (2015) theory of variation.

The nature of Level 1 boundaries may be arbitrary in that there are infinitely many ways experience may be bounded, but they nonetheless define the premises from which necessary implications may be derived. It is here that the illusion of acquire-able "things" emerges. When we give differences a name, they appear as objects—as a "this" instead of a "this-as-opposed-to-that" or the conceptual "space-between-this-and-that". But if we lose sight of the difference that the name points to, we put ourselves and our students in a position from which no further insight is possible.

We take as our starting point exponents *as a special case of repeated multiplication in which all multiplicands are the same*. Prior experience with varied interpretations of multiplication will be assumed (see Fig. 12.7) as the starting point from which further differentiation may be prompted.

The *new* features to be developed through the full Level 1–5 sequence are (a) the extension from three multiplicands to an infinite number of addends, which has different implications for different interpretations of multiplication (Davis 2015), but which we wish to generalize to numerical laws, (b) the repetitive nature of multiplication when working with exponents, and (c) the mathematical notation used to

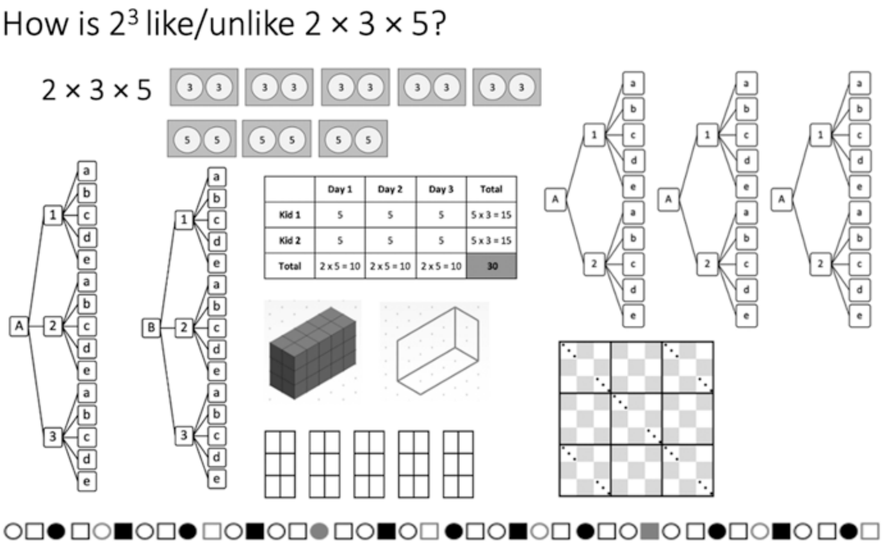

Fig. 12.7 How is 2^3 like/unlike $2 \times 3 \times 5$?

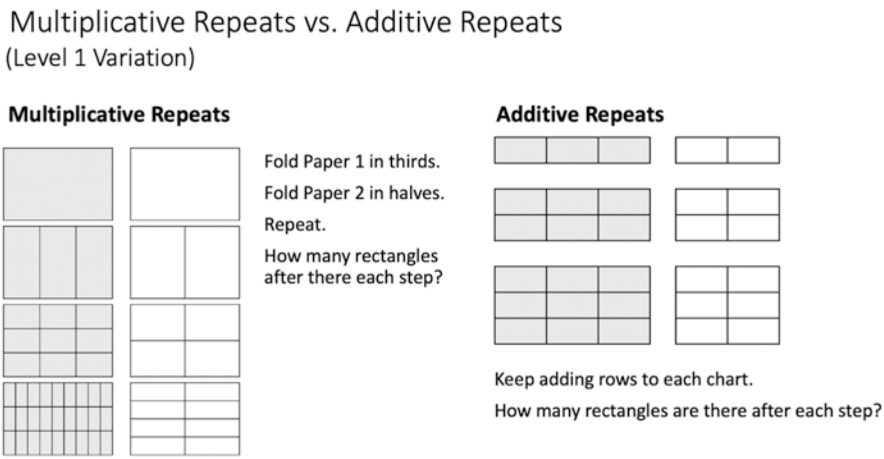

Fig. 12.8 Level 1 variation: What is repeated multiplication?

describe varied configurations of repeated multiplication. The decision to bound the multiplicands in such a way that they all match and the use of exponential notation to describe the possibilities that emerge in that space is arbitrary (Hewitt 1999), but clear contrasts are useful in prompting attention to these boundaries (see Figs. 12.8 and 12.9). We define what exponents *are* through contrast through what they are *not*, then generalize to less standard or more complex cases. Level 1 can often be characterized in terms of "yes-no-also," as in Fig. 12.9.

What *Are* Exponents?
(Level 1 Variation)

Exponents define *number of multiplicative repeats (starting at 1).*

YES

$2^0 = 1$

$2^1 = 2$

$2^3 = 2 \times 2 \times 2 = 8$

$2^5 = 2 \times 2 \times 2 \times 2 \times 2 = 32$

$10^0 = 1$

$10^1 = 10$

$10^2 = 10 \times 10 = 100$

$10^5 = 10 \times 10 \times 10 \times 10 \times 10 = 10{,}000$

NO

$2^5 \neq 2 \times 5$

$25 \neq 10$

$10^5 \neq 10 \times 5$

$10{,}000 \neq 50$

$2^5 \neq 5^2$

$32 \neq 25$

ALSO: $(?)^3 = ? \times ? \times ?$

$(\mathbf{6+2})^3 = (6+2) \times (6+2) \times (6+2) = 8 \times 8 \times 8 = 512$

$(\mathbf{6-2})^3 = (6-2) \times (6-2) \times (6-2) = 4 \times 4 \times 4 = 64$

$(\mathbf{6 \times 2})^3 = (6 \times 2) \times (6 \times 2) \times (6 \times 2) = 12 \times 12 \times 12 = 1728$

$(\mathbf{6 \div 2})^3 = (6 \div 2) \times (6 \div 2) \times (6 \div 2) = 3 \times 3 \times 3 = 27$

$((\mathbf{6 \div 2})^2)^3$
$= (6 \div 2)^2 \times (6 \div 2)^2 \times (6 \div 2)^2$
$= 3^2 \times 3^2 \times 3^2$
$= 9 \times 9 \times 9$
$= 729$

$((\mathbf{6 \div 2})^3)^2$
$= (6 \div 2)^3 \times (6 \div 2)^3$
$= 3^3 \times 3^3$
$= 27 \times 27$
$= 729$

Fig. 12.9 Level 1 variation: What are exponents?

Varying Problem Features
Level 2 Variation: *Which column prompts* more effectively?*

How Alike? (Induction)

What do all have in common?

a) $3^5 \times 3^2$
$= (3 \times 3 \times 3 \times 3 \times 3) \times (3 \times 3)$
$= 3^7$

b) $4^6 \times 4^2$
$= (4 \times 4 \times 4 \times 4 \times 4 \times 4) \times (4 \times 4)$
$= 4^8$

c) $7^2 \times 7^3$
$(7 \times 7) \times (7 \times 7 \times 7)$
$= 7^5$

d) $6^{12} \times 6^3$
$(6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6) \times (6 \times 6 \times 6)$
$= 6^{15}$

How Different? (Generalization)

What happens if I change…?

a) $3^4 \times 3^3$
$(3 \times 3 \times 3 \times 3) \times (3 \times 3 \times 3)$
$= 3^7$

b) $\mathbf{8}^4 \times \mathbf{8}^3$
$(\mathbf{8} \times \mathbf{8} \times \mathbf{8} \times \mathbf{8}) \times (\mathbf{8} \times \mathbf{8} \times \mathbf{8})$
$= \mathbf{8}^7$

c) $8^4 \times 8^{\mathbf{5}}$
$(8 \times 8 \times 8 \times 8) \times (8 \times 8 \times 8 \times \mathbf{8} \times \mathbf{8})$
$= 8^{\mathbf{9}}$

d) $8^4 \times \mathbf{9}^5$
$(8 \times 8 \times 8 \times 8) \times (\mathbf{9} \times \mathbf{9} \times \mathbf{9} \times \mathbf{9} \times \mathbf{9})$
$= 8^4 \times \mathbf{9}^5$ (can't be combined)

*Note this is just the first part of the prompt—learners would then *engage in a similar set of practice prompts.*

Fig. 12.10 Level 2 variation: Varying base, exponent

Level 2: Varying Features

Having thus defined the space in which we wish to work, we may now focus on variation of those apparent "things" or features. This puts us at Level 2, differences between differences, which has much in common with Marton's generalization and with the first type of procedural variations (PV1) described by Lai and Murray (2014); i.e., variation of problem conditions. This is highlighted in Fig. 12.10, which also attempts to clarify the difference between an inductive approach to varying problem conditions (not recommended) and an approach focused on difference (recommended).

Already in Fig. 12.9, variations of "yes" and "no" were offered, but the primary purpose of given contrasts was to identify relevant features (base, exponent) and to highlight the rules for interpreting exponential notation. At Level 2, we start to explore the *implications* of varying features identified at Level 1. In short, from Level 1 to Level 2, the goal shifts from defining boundaries to exploring implications of change. Importantly, changing one feature (A) has a resultant change on another (B), and the focus of attention shifts to this *relationship*, which might be seen as a sort of *difference between A and B*, and which we will call C. It's important that first one feature varies, then another, and then both together—Marton's *fusion*. Changing which feature is known and which is unknown can deepen understanding of the relationship between identified variables.

On the left side of Fig. 12.10, several examples are given in which learners must add exponents to get an answer. However, each question uses new bases and new exponents, which can make it harder to see the impact of change. In such cases, it is easy to fall into the trap of the teacher asking learners to "guess what is in my mind" (Mason 2010). While what is the same may seem obvious to the teacher, there are, in fact, a variety of features that are the same, and it's not always easy for learners to zone in on the intended one. In cases like this, teachers typically end up *giving* the rule, then ask learners to apply it in multiple cases. In so doing, it essentially gets reduced to a procedure rather than a generalized relationship.

The examples on the right are also cases where learners must add exponents, but now only one feature changes in each question, which makes it easier both to identify the intended feature *and* to see the *impact* of each change. Whether we are varying *within a particular law* or *between multiple laws*, we can limit variation to all but one feature, then observe the impact of changing that feature. Note that examples we've highlighted primarily focus on the "change" aspect of "change and choice." Through careful questioning as varied examples are offered and through appropriate follow-up tasks, it is also important that the teacher require learners to *make* relevant distinctions.

Level 3: Varying Relationships

If relationships *within a particular exponent law* are the focus, then it makes sense to vary features of that law—one feature at a time—and to observe the effect of doing so. By highlighting those changes—and their impact—the focus of attention shifts to relationships. However, distinguishing *among different exponent laws* should also become a focus fairly quickly (see Fig. 12.11). To do so, it is helpful to contrast the laws themselves while holding as many features constant as possible. In the set on the right, the bases and exponents are kept constant, while the operations change. Again, instead of doing several examples that all require adding exponents

Contrasting Relationships

Level 3 Variation: From difference within to difference between

Difference Within (Level 2)

a) $3^4 \times 3^3$
$= (3 \times 3 \times 3 \times 3) \times (3 \times 3 \times 3)$
$= 3 \times 3 \times 3 \times 3 \quad \times \quad 3 \times 3 \times 3$
$= 3^7$

b) $\mathbf{8}^4 \times \mathbf{8}^3$
$= \mathbf{8 \times 8 \times 8 \times 8} \quad \times \quad \mathbf{8 \times 8 \times 8}$
$= \mathbf{8}^7$

c) $8^4 \times 8^5$
$= 8 \times 8 \times 8 \times 8 \quad \times \quad 8 \times 8 \times 8 \times \mathbf{8 \times 8}$
$= 8^{\mathbf{9}}$

d) $8^4 \times \mathbf{9}^5$
$= 8 \times 8 \times 8 \times 8 \quad \times \quad \mathbf{9 \times 9 \times 9 \times 9 \times 9}$
$= 8^4 \times \mathbf{9}^5$ (can't be reduced)

Difference Between (Level 3)

a) $3^2 \times 3^6$
$= 3 \times 3 \quad \times \quad 3 \times 3 \times 3 \times 3 \times 3 \times 3$
$= 3^{2+6} = 3^8$

b) $(\mathbf{3^2})^6$
$= (\mathbf{3 \times 3})^6$
$= \mathbf{3 \times 3} \quad \times \quad \mathbf{3 \times 3} \quad \times \quad \mathbf{3 \times 3} \quad \times \quad \mathbf{3 \times 3} \quad \times \quad \mathbf{3 \times 3} \quad \times \quad \mathbf{3 \times 3}$
$= 3^{6x2} = 3^{12}$

c) $3^6 \div 3^2 = 3^4$
$= \dfrac{3 \times 3 \quad \times \quad 3 \times 3 \quad \times \quad \cancel{3 \times 3}}{\cancel{3 \times 3}} = 3^{6-2} = 3^4$

d) $3^6 \div 3^4 = 3^2$
$= \dfrac{3 \times 3 \quad \times \quad \cancel{3 \times 3} \quad \times \quad \cancel{3 \times 3}}{\cancel{3 \times 3 \times 3 \times 3}} = 3^{6-4} = 3^2$

Fig. 12.11 Level 3 variation: Contrasting exponent laws

and then identifying what they have in common, here the focus is on impact of change.

This pattern or variation may still be seen in terms of Lai and Murray's PV1 (changing features), but there are also elements of PV2, or changing strategies. In our case, this doesn't show up in the sense of multiple strategies to solve the same problem but through the manner in which both Level 3 and PV2 prompt to contrasting relationships. In fact, Level 3 may be seen as a version of PV2 that varies the relationship against a constant background instead of varying the representation or context against a constant relationship. This isn't to say that there isn't value in comparing carefully selected multiple strategies (Durkin et al. 2017), but for reasons that we explain a bit later, doing so is better described by Level 5 (integration) in our scheme.

Level 4: Abstraction

At Level 4, the relationships among relationships themselves become the focus of attention through contrast with other situations that partially share explanatory structure; meaning may move in both directions, but it's helpful when at least one situation is clearly understood. The example we offer in Fig. 12.12 may seem like an application rather than an abstraction, but the point is that it offers a space where the structure of relevant relationships may be contrasted in ways that allow transfer of meaning.

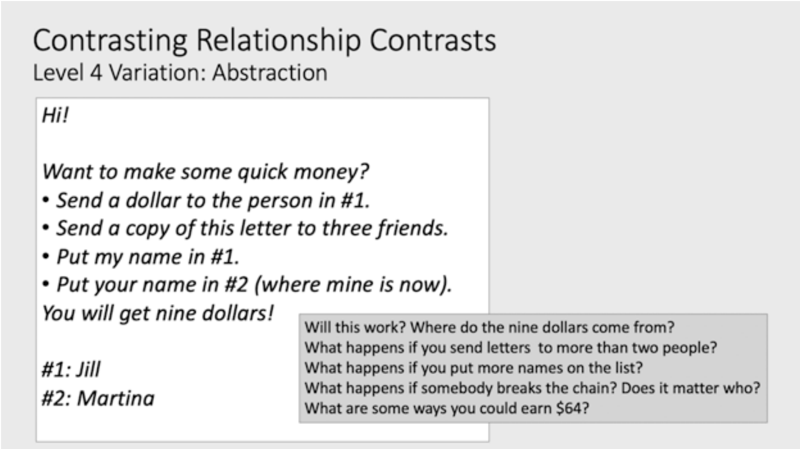

Fig. 12.12 Level 4 variation—Contrasting relationship contrasts: Making sense of exponential structure

The variables in the chain letter (i.e., number of people you send the letter to, number of people on the list, number of people who break the chain) could be seen in terms of variable bases and exponents, but recognizing this as a possibility and considering the appropriateness of transfer is key to making sense of the task. Consideration of combination locks with variable numbers of both numbers (or other symbols) on the lock and numbers (or other symbols) in the combination would serve a similar purpose. The variations described here are partially consistent with Lai and Murray's PV3, which involves "multiple applications of a method by applying the same method to a group of similar problems" (p. 8). Again, however, our emphasis is not on whether the task offers an *application* but on whether the similarity between problems affords transfer of explanatory structure (which is consistent with the examples offered by Lai and Murray).

Lest it seem that we are contradicting ourselves in recommending another strategy that explicitly focuses on identifying similarity, note that mapping an analogy differs from induction in essential ways. *Induction* inappropriately focuses on perception of similarity in that it requires learners to find commonalities among features they have not yet discerned. Earlier, we highlighted how generalization focuses on similarity between previously discerned *features*; here, we highlight how *abstraction* focuses on similarity between previously discerned *relationships*.

Level 5: Integration

Level 5 focuses on the use of tasks that require the integration of seemingly unrelated concepts that become enmeshed in the "same story" and thereby assume relevance to one another. In doing so, Level 5 incorporates all levels of the Level 1–4 hierarchy without adding another layer of differences among differences; it does so by bringing together ideas previously formed through their own progressions. We distinguish integration from abstraction in that where abstraction focuses on the transfer of meaning between two at least partially familiar situations, integration requires the combining of familiar ideas to solve a problem. Integration, then, does not fit into the same hierarchical structure described in Levels 1–4.

The rope metaphor we used to describe the importance of integrating prior knowledge that is well-understood—of taking care not to ask students not to braid the strands of a rope at the same time that they're asked to braid those strands into a rope—is key to Level 5. If raveled appropriately, the "How Many Factors" task we introduced earlier (Fig. 12.5) may be seen as a Level 5 task: It *requires* and therefore *integrates* ideas developed in multiple Level 1–4 progressions, including the one pertaining to exponents developed here. Thus, it might be considered a Level 5 task in a variety of progressions, depending on the order in which topics were introduced. The key point is that each of the components *has* been previously developed before we ask learners to integrate them.

Again, it may seem that integration is mere application. Both abstraction and integration, may (but need not) overlap with applications, but the notion of application doesn't seem to be a particularly useful distinction when considering how a problem or task set prompts to meaning. Similarly, the use of multiple strategies to solve a problem (and considering how they're related) may focus either on varying relationships (Level 3) or on integrating diverse ideas, which we argue is an important distinction.

A Brief Note on Problem Solving

Various notions of problem solving assume relevance in different places within the hierarchy we've developed here, particularly those that focus on general heuristics and those that emphasize transfer to novel situations. A clear focus on the ongoing structuring of knowledge engages learners directly with the sorts of ideas typically highlighted in lists of problem solving strategies, and the dynamic structures thus developed lend themselves metaphoric transfer. Further, Levels 1–3 focus on what is sometimes referred to as "working systematically." Here, working systematically is focused neither on procedural steps with clear worked examples nor "rich tasks" with a focus on "mathematical process."

Teaching with variation *models* working systematically (with structured variation) and weaves together mathematical ideas that *support* working systematically (e.g., identifying combinations, graphical representation, algebraic representation). In other words, ideas are raveled into co-amplifying ideas that serve *each other*. As learners become more familiar with using structured variation and more aware of dimensions that are vary-able, they can take greater responsibility for creating the variations that prompt to new meaning. Working systematically typically gets short-changed in one of two ways: (1) When it's seen as *only* a process for approaching other content, it doesn't get adequately raveled in its own right and (2) when it's seen *only* as an isolated body of content, the important role it *does* play in making sense of other content gets overlooked. Both matter—and both are borne of the same artificial process-content dichotomy.

Levels 4 and 5 address problem solving as application, but application is divided into two categories that differ in terms of the ways they support structuring meaning: Level 4 is about recognizing structural similarity among diverse problems, while Level 5 is requires decomposing complex problems into manageable subproblems (here "manageable" includes requisite prior knowledge). In other words, Levels 4 and 5 draw from and elaborate the structured knowledge developed in Levels 1–3. Various conceptions of modeling might similarly be categorized according to their role in structuring meaning.

While various discourses on mathematical problem solving (English and Gainsburg 2016; Liljedahl et al. 2016) acknowledge the role of prior knowledge, they focus less on the long-term structuring of that knowledge than on the immediate actions (or non-actions) taken in the hopes of calling forth or generating a fruitful combination of ideas relevant to a particular problem situation. While problem-solving heuristics may support mathematical creativity, discovery, and invention, they can easily lend themselves to an air of mystique that perpetuates the myth that mathematics is primarily the realm of those with a particular type of ability or even genius. We hope that our emphasis on long-term structuring helps create rich ground from which all learners may share in the powerful a-has that accompany moments of illumination and insight.

Summary

Most reports of variation pedagogy focus on a short-term ravel, likely with the assumption that what is well-integrated locally will also be well-integrated on a broader scale. As we observe variation pedagogy being taken up outside of China, this does not seem to be a well-grounded assumption. Collections of isolated lessons, even when well-varied internally (i.e., even when the focus is on mathematical structure and not merely on procedures), do little to prompt to the broader integrated structure of mathematics, and most curricula and resources are grossly insufficient in supporting such coherence (Bajaj 2013).

It is a monumental undertaking to build an effective sequence of effectively varied lessons. Even when thoughtfully designed, such lessons inevitably change when they meet students—not just to suit the idiosyncrasies of individual students, but in ways that gradually become more consistent with the phenomenographic space of possibility defined by the affordances of human perception. It is not reasonable to expect individual teachers to take on the task of re-inventing and refining effective long-term mathematical sequences. Even when teachers have sufficient pedagogical content knowledge relevant to their grade level, there is much that can be offered in the form of a well-raveled resource that takes into account both the structure of mathematics, common patterns of interaction between learners and mathematics, and long-term coherence between grades. With this in place, teachers are simultaneously supported in making sense of the long-term ravel and freed to attend to the fine-grained variations relevant to their moment-by-moment interactions with students.

Conclusion

As both teachers and researchers, we must confess to a frustration with the field of education. At times it feels as though there is no other domain of human engagement that is more resistant to well-structured theory (e.g., exposing the metaphoric substrate of entrenched activity) and validated evidence (e.g., year-over-year improvements in learner engagement and understanding). When presented with such ideas and evidence, more often than not, the system finds ways to reject or minimize it—often by characterizing a new insight as reflective of the other "camp" in whatever skirmish is happening at the moment.

It's thus that we have experienced criticisms and rejections from teachers and policy makers positioned at both poles in the Math Wars. For instance, Traditionalists balk at the assertions that *all* learners can become adept at mathematics, that gaps in understanding are attributable to missed noticings, and that perceived differences in learner ability have more to do with flawed pedagogy than flawed learners. On the other side, Reformists have a strong tendency to see a well-deconstructed concept in terms of the much-hated step-based approach to teaching rather than an equity-informed noticing approach. Absences are another favorite focus of their criticisms. Limited group work, few open-ended problems, and no heed to personalized strategies—notwithstanding that the evidence supporting such emphases is dubious at best—are lobbed as reasons to reject the model entirely. We actually take these rejections by both staunch Traditionalists and staunch Reformists as positive indicators, emboldened by Dewey's (1990: 9) observation regarding oppositional thinking, noted earlier: "We don't solve them: we get over them."

Part of that "getting over" is hinged to rethinking the relationship between teachers and resources. As we hope is evident in the preceding discussion, a well-structured inquiry involves high levels of knowledge and extensive effort. Each of our lessons has pulled in the expertise of mathematicians, logicians, teachers, and educational researchers. Flatly stated, there is no way that solitary teachers in isolated classrooms might be reasonably expected to design such lessons on their own.

It is thus that we frame fitted-to-the-Information-Age approach to mathematics teaching in terms of a partnership in which each aspect is associated with differentiated obligations. Principal responsibility for the Ravel—that is, discerning the critical discernments involved in a concept, appreciating their relationships to one another, and so on—sits with the resource. Responsibility for the Prompt is more shared, with the resource providing suggestions and that teacher selecting and adapting those suggestions according to knowledge of experiences, established competencies, and interests of those present. The contingencies associated with Interpreting means that that element is almost entirely the responsibility of the teacher, and Deciding what to do next is a shared responsibility that sits across the teacher's knowledge of what's happening and the resource's advice on what might happen next. Conceived as a partnership, the RaPID model is neither a step-following (Traditionalist) script nor an open-ended (Reformist) exploration.

To state this point more emphatically, we see the next moment in the necessary evolution of school mathematics in the Information Age to be about a much-expanded and formalized partnership between teachers and resources, each having obligations to the other.

That suggestion is heresy within much of the current educational establishment. It strikes against two principles that are held by Traditionalist and Reformist alike: firstly, a conviction on the sanctity of teacher autonomy and, secondly, a belief that the best response to learner difference is differentiated experience. We question those ideals. In an era of massive connectivity (in which there can be genuine, mutually beneficial influences between teachers and resource developers) and better understandings of knowledge and cognition (that point to flawed assumptions in differentiated models of instruction), new possibilities for school mathematics are not just afforded, they are required.

References

Bajaj, V. (2013). Q and A with Liping Ma. *The New York Times*. From https://www.nytimes.com/2013/12/18/opinion/q-a-with-liping-ma.html

Bateson, G. (2002). *Mind and nature: A necessary unity.* Cresskill: Hampton Press. (Original work published 1979).

Chazan, D. and Ball, D. (1999). Beyond being told not to tell. *For the Learning of Mathematics* 19: 2–10.

Davis, B. (2015). Exponentiation: A new basic? *Mathematics Teaching in the Middle School* 21: 34–41.

Dewey, J. (1910). *The influence of Darwin on philosophy and other essays*. New York: Henry Holt.

Dienes, Z. (1960). *Building up mathematics*. New York: Hutchinson.

Durkin, K., Star, J., and Rittle-Johnson, B. (2017). Using comparison of multiple strategies in the mathematics classroom: lessons learned and next steps. *ZDM Mathematics* 49: 585–597.

English L. D. and Gainsburg, J, (2016). Problem solving in a 21st century mathematics curriculum. In: L. D. English and D. Kirshner (eds.), *Handbook of international research in mathematics education* (3rd edition), 313–335. New York: Routledge.

Foote, R. (2007) Mathematics and complex systems. *Science* 318: 410–412.

Gu, F., Huang, R., and Gu, L. (2017). Theory and development of teaching through variation in mathematics in China. In: R. Huang and Y. Li (eds.), *Teaching and learning mathematics through variation: Confucian heritage meets Western theories,* 13–41. Boston: Sense Publishers.

Gu, L., Huang, R., and Marton, F. (2004). Teaching with variation: A Chinese way of promoting effective mathematics learning. In: L. Fan, N.-Y. Wong, J. Cai, and S. Li (eds.), *How Chinese learn mathematics: Perspectives from insiders*, 309–347. Hackensack: World Scientific.

Hewitt, D. (1992). Train-spotters' paradise. *Mathematics Teaching* 140: 6–8.

Hewitt, D. (1999). Arbitrary and necessary: Part 1: A way of viewing the mathematics curriculum. *For the Learning of Mathematics* 19: 2–51.

Hewitt, D. (2001a). Arbitrary and necessary: Part 2: Assisting memory. *For the Learning of Mathematics* 21: 44–51.

Hewitt, D. (2001b). Arbitrary and necessary: Part 3: Educating awareness. *For the Learning of Mathematics* 21: 37–49.

Huang, R., Barlow, A., and Prince, K. (2016). The same tasks, different learning opportunities: An analysis of two exemplary lessons in China and the U.S. from a perspective of variation. *The Journal of Mathematical Behavior* 41: 141–158.

Huang, R. and Leung, K. S. F. (2004). Cracking the paradox of Chinese learners. In: L. Fan, N.-Y. Wong, J. Cai, and S. Li (eds.), *How Chinese learn mathematics: Perspectives from insiders*, 348–381. Hackensack: World Scientific.

Jianhua, L. I. (2004). Thorough understanding of the textbook—A significant feature of Chinese teacher manuals. In: L. Fan, N.-Y. Wong, J. Cai, and S. Li (eds.), *How Chinese learn mathematics: Perspectives from insiders*, 262–281. Hackensack: World Scientific.

Kullberg, A., Runesson Kempe, U., and Marton, F. (2017). What is made possible to learn with the variation theory of learning in teaching mathematics? *ZDM Mathematics Education* 49: 559–569.

Lakoff, G., and Núñez, R. (2000). *Where mathematics comes from*. New York: Basic Books.

Lai, M. Y. and Murray, S. (2014). Teaching with procedural variation: A Chinese way of promoting understanding of mathematics. *International Journal for Mathematics Teaching and Learning.* from http://www.cimt.org.uk/journal/lai.pdf.

Liljedahl, P., Santos-Trigo M., Malaspina, U., and Bruder, R. (2016). *Problem solving in mathematics education.* Chaim: Springer Nature.

Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahwah: Lawrence Erlbaum.

Marton, F. and Booth, S. (1997). *Learning and awareness.* Mahwah: Lawrence Erlbaum.

Marton, F. and Tsui, A. B. M. (2004). *Classroom discourse and the space of learning*. Mahwah: Lawrence Erlbaum.

Marton, F. (2015). *Necessary conditions of learning*. New York: Routledge.

Mason, J. (2010). Effective questioning and responding in the mathematics classroom. from http://mcs.open.ac.uk/jhm3/Selected%20Publications/Effective%20Questioning%20and%20Responding.pdf

Mason, J. and Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics* 15: 277–289.

Mason, J., Burton, L., and Stacey, K. (2010). *Thinking mathematically* (2nd edition, New York: Prentice Hall.

Math Minds. (2020). *Math minds online course*. https://www.structuringinquiry.com.

Preciado-Babb, A. P., Metz, M., Davis, B., and Sabbaghan, S. (2020). Transcending contemporary obsessions: The development of a model for teacher professional development. In: S. Linares and O. Chapman (eds.), *International handbook of mathematics teacher education*. 361–390. Boston: Sense Publishers.

Ong, W. J. (1982). *Orality and literacy: The technologizing of the word*. New York: Routledge.

Organization for Economic Cooperation and Development. (2010). *Recognising non-formal and informal learning: Outcomes, policies and practices.* http://www.oecd.org/education/innovation-education/recognisingnon-formalandinformallearningoutcomespoliciesandpractices.htm

Pang, M.F., Marton, F., Bao, J., and Ki, W.W. (2016). Teaching to add three-digit numbers in Hong Kong and Shanghai: Illustration of differences in the systematic use of variation and invariance. *ZDM Mathematics Education* 48: 455–470.

Runesson, U. (2005). Beyond discourse and interaction. Variation: A critical aspect for teaching and learning mathematics. *Cambridge Journal of Education* 35: 69–87.

Schoenfeld, A. H. (2004). The math wars. *Educational Policy* 18: 253–286.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher* 27: 4–13.

Skemp, R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77: 20–26.

Sun, X. (2011). "Variation problems" and their roles in the topic of fraction division in Chinese mathematics textbook examples. *Educational Studies in Mathematics* 76: 65–85.

Watson, A. (2017). Pedagogy of variations: Synthesis of various notions of variation pedagogy. In: R. Huang and Y. Li (eds.), *Teaching and learning mathematics through variation: Confucian heritage meets Western theories,* 85–103. Boston: Sense Publishers.

Watson, A. and Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Mahwah: Lawrence Erlbaum.

Watson, A. and Mason, J. (2006). Seeing an exercise as a single mathematical object: Using variation to structure sense-making. *Mathematical Thinking and Learning* 8: 91–111.

Wong, N. Y. (2008). Confucian heritage culture learner's phenomenon: from "exploring the middle zone" to "constructing a bridge." *ZDM Mathematics Education* 40: 973–981.