1 Watson-Education's ambition: *Competency-based* personalized learning.

Some critics of public education argue that its dominant operating paradigm is an outmoded industrial- age construct [Bramante2012]. It implicitly assumes a strong, direct correlation between what students know and the amount of time spent sitting in a Education. Consider use of terms like "semester hours" or "credit hours". In contrast, "Competency-based education is a term that describes learning progressions based on mastery of content rather than passage of time" [NCSL2017]. We emphasize, here, *competency-based*. The U.S. Department of Education [USDoEd] places the two in the same breath. The emphasis on competency occurs elsewhere in education literature. Examples include [Bramante2012], [Rickabaugh2016], [Tomlinson2014], [Marzano2008], [Drake2004], and [Wiggins2005].

Competency-based learning stands in contrast to a content- or style-focused approach. Visual, Aural, Read/Write, Kinesthetic (VARK) [Flemming2001] represents one popular style-focused approach to learning personalization. Watson Education allows teachers record their opinions about learners' VARK characteristics. The education community's view however of the veracity of the science underlying VARK falls short of a unanimous consensus (e.g., [Khazan2018]).

Competency-based learning in not new. References to competency-based learning appear in research literature as early as the 1970s (e.g., [Murray1976]). Loosely-related to competency-based learning, self-paced learning goes back at least as far. Predecessors to contemporary SRA Math laboratories [SRA] were used for self-paced strategies in elementary and middle schools beginning in the mid-1970s.

Truly personalizing learning or differentiating instruction requires Education teachers process substantial quantities of information. Continuous understanding of each student's state of knowledge at a granular level represents the starting point. It has only become practical given recent information-management technologies. These include data organization, retrieval, presentation, and analytics. IBM seeks to bring its corporate competencies to bear to make personalization of learning practical at scale.

1.1 Watson-Education's vision for personalization of learning and differentiation of instruction.

IBM's strategy for the personalization of learning seeks to answer three questions about individual learners' trajectories through their educational curriculum. Figure 1.1 illustrates conceptually. "Know me" is the first question. Conceptually, Watson-Education — IBM's cloud-based personalized-learning offering — groups essential elements of information about learners into three

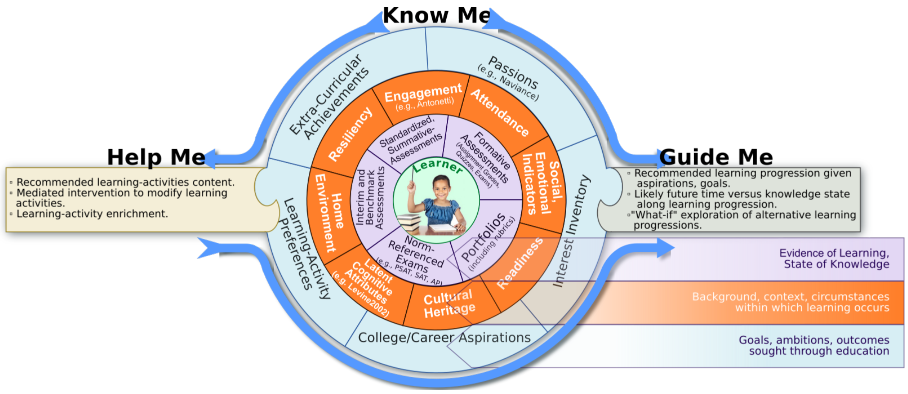

Figure 1: Watson Education assists educators in personalization of learning using a "Know-me, Help-me, Guide-me" framework.

layers. The inner layer — lavender-colored in Figure 1.1 — pertains to the learner's state of knowledge.

The orange-colored ring in the figure pertains to situational factors. These may lead to barriers to learning that teachers may need to remove or to work around. Watson-Education aspires to surface them early to give teachers the opportunity to be proactive.

The learner attributes in the orange-colored are often referred to as "non-cognitive" factors of learning. The Univerisity of Chicago Consortium on Chicago School Research (UCCSR) surveys the extensive literature on non-cognitive affects' influence on learning outcomes [Farrington2012]. Watson Education incorproates provides Education teachers with views on these factors based on two sources. First, peer education-technology systems — such as Student Information Systems (SISs) — provide data extracts that are ingested into Watson.

Additionally, teachers record varieties *ad hoc* observations in the Watson-Education application. These *ad hoc* observations document evidence of learners' extent of academic and social engagement in their educational environment. These observations address many of the factors in the orange-colored ring in Figure 1.1.

The blue-colored ring represents an aspirational view of the learner. It provides a view of an answer to the student's question, "What do I want to be when I grow up?" The content and trajectory of each learner's curriculum should be influenced by the aspirational view.

Next, Watson-Education seeks to personalize learning by answering the "Help Me". This pertains to the day to day work of helping learners get through the curriculum. Watson-Education makes easily accessible supplementary instructional materials. Having alternative content at their fingertips helps teachers guide students around rough spots.

Finally, the "Guide Me" question pertains to the aspirational view. At the time of this writing, this is an aspirational feature of the offering. Students should have a differentiated — if not personalized — curriculum. This curriculum must address high-school graduation requirements, the present scope of the offering. For example, a learner aspiring to a Science, Technology, Engineering, or Mathematics (STEM) career should follow a curriculum distinguished from another pursuing the performing arts.

1.2 Education operational frameworks underlying Watson-Education functionality.

IBM's design of Watson Education seeks to remove specific barriers to the adoption of competency-based personalized learning at scale. The guiding philosophy is described by Edelmen [Edelmen2015] and Christensen [Christensen2016]. IBM thought leaders were also influenced by Manning, *et al* [Manning2012]. IBM seeks to position its offering on the leading edge of educational practice.

IBM does not view the personalization of learning as purely a technology play. Contemporary business research indicates that, "Strategy, not technology, drives digital transformation" [Kane2015]. Personalization of learning at-scale requires — for reasons mentioned above — digital transformation within education organizations.

The principle barriers to wide-scale adoption of competency-based education are both operational and informational. Abandoning a uniform, assembly-line approach to Education instruction requires teachers to operate their Educations in a fundamentally different manner. This operating approach moreover requires access to continuously updated information.

Traditional, time-based education — at its worst — is a linear, static framework. Lessons are planned and delivered. Learners are assessed. The class moves on. Students who miss out on a key idea are often largely left to their own devices to catch up. A missed concept in one course unit can dow the seeds for subsequent learning struggles lasting potentially for years. Frustrated students disengage or even drop out for these reasons.

1.2.1 Fundamental underlying premise: The progressive nature of learning.

Watson Education operates on the fundamental premise that knowledge is cumulative. This theme recurs throughout the education literature. Explicitly, the Math learning-standards document for the Common Core State Standards (CCSS) states, "What students can learn at any particular grade level depends upon what they have learned before. Ideally then, each standard in this document might have been phrased in the form, 'Students who already know ... should next come to learn ....'" [CCSS-Math, p. 5]

Similarly, Almond, *et al*, observe, "The student's probability of knowing the skill at the end of the course will depend both on the student's skill level at the beginning of the course, and what kind of instruction the student receives" [Almond2015, p. 98]. Almond speaks of instruction at the course level. One obviously can think at a more-granular level, down to the learning activity.

Stevens, *et al*, drill down to this level, "Learning progressions describe how students gain more expertise within a discipline over a period of time. They represent not only how knowledge and understanding develops, but also predict how the knowledge builds over time. Thus, the focus is not limited on the end-product knowledge as characterized by summative assessment, but on how students' ideas build upon other ideas" [Stevens2007]. The idea of "ideas building upon other ideas" lies at the root of the operating premise on which Watson-Education is based.

The CCSS-Math document observed at the time of its writing that "...existing education research cannot specify all such learning pathways" [CCSS-Math, p. 5]. Nonetheless, identifying progressions of learning subsequently became an area of active research. A group of CCSS math-standards authors wrote a set of working documents [Askey] describing progressions of CCSS Math standards. Team member Jason Zimba undertook construction of a graph of the pre-secondary math standards [Zimba2012]. This formed the basis of a "Coherence Map" offering by nonprofit education-technology organization Student Achievement Partners [SAP]. The University of Kansas developed a more-extensive offering, Dynamic Learning Maps [DLM].

Many sets of learning standards developed subsequent to the CCSS contain built-in learning progressions. The Next-Generation Science Standards (NGSS) contains an appendix describing progressions of disciplinary core ideas [NGSSApdxE]. The State of New York plans adoption of a new set of Mathematics learning standards containing a coherence "map" within the standards document itself [NYNGMath]. The Texas Education Agency published vertical alignments for Math [TXGateway], and for English, Language Arts, and Reading [TEA2017] for its state learning standards. IBM engaged nationally recognized curriculum export Dr. Karen Hess to develop CCSS ELAR vertical alignments based on previous work [Hess2011].

Recognition of the progressive nature of learning — based on learning progressions is slowly moving into the mainstream of education practice. IBM's Watson-Education stakes out — in the spirit of "competing on customer journeys" [Edelmen2015] — a position on the leading edge of this stage in the evolution of education practice.

It must be noted however that not all learning is inherently progressive. Psychologists distinguish between knowledge types [tenBerge1999, Yilmaz2012].

Procedural knowledge tends to be progressive in nature. Declarative knowledge tends to be fact- or rule-based [Wiggins2005, p. 341].

Knowledge in mathematics and the exact-science domains tends to be procedural. Knowledge in English, Language Arts, and Reading (ELAR) and Social Studies and Humanities tends to be declarative. Watson-Education's progression-focused Mastery works better for procedural-knowledge domains than for the declarative.

1.2.2 Central question for competency-based personalized learning: *What is each learner's state of knowledge*?

Deriving business value from any artificial intelligence (AI) or analytics solution begins with well-posed questions [LaValle2011, Satell2018]. Watson-Education mastery focuses on *a narrow question*. This question is, "What do we believe the state of each learner's knowledge with respect to each learning standard given all available evidence of learning?"

Educators ask this question about their students in both short-term and long-term contexts. The short-term form of the question asks, "Is each student prepared for the instructional material I am now going to present?" This question directly addresses Almond and Mislevey's statement above [Almond2015, p. 98] about the progressive nature of learning.

Educators also ask questions about their students' state of knowledge with respect the broader scope of College and Career Readiness, and for accountability systems — "high-stakes" standardized exams. This broader state-of-knowledge question measures the voilet-colored ring in Figure 1.1 with respect to requirements in the blue-colored ring. Watson Education — at the time of this writing — only attempts to answer this question in to the extent of readiness for high-stakes accountability exams and high-school graduation.

Watson Eduction answers these questions through a rudimentary *Diagnostic Cognitive Model* (DCM). DCMs belong to a family of probabilistic models called *Latent Variable Models* (LVMs) ([Loehlin2004], [Rupp2010], [Leighton2007]). LVMs are *latent* because they make inferences about conditions that are not directly observable. We elaborate at length in section 2.

Figure 2 illustrates more-detailed interpretations of the answers the short-term question about students' readiness for the instructional content presented next. Watson-Education first exhibits the body of evidence available on which to form beliefs about individual learners' knowledge states. These are the inputs to the cognitive-diagnostic model. The Figure illustrates the use of a Bayesian Network to answer Watson-Education's central questions. We elaborate on this approach in subsequent sections.

Answers to subsequent questions are based on Watson Education's fundamental premise about the progression of learning. Watson Education retro-

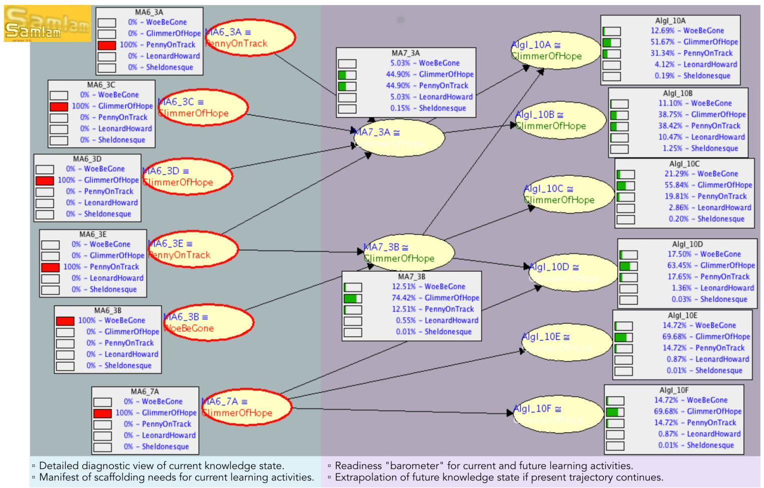

Figure 2: Watson Education's Mastery uses a Diagnostic Cognitive Model (DCM) (e.g., [Rupp2010]) to answer key questions about each individual learner's state of knowledge and readiness to engage instructional content planned next. (Visualization by [SAMIAM].)

spectively diagnoses likely explanations for weakness when given mixed results in terms of measured learning standards. It answers the question, "What scaffolding — supplementary or remedial content — should be included in the instruction in order to address weaknesses?"

The third question is forward-looking. Watson Education answers the question, "How reader is each learner to engage the instructional content I am about to present?" IBM draws sharp distinction between the tempting, alternative question, "How well will my learner perform on the instructional material I am about to present?" As stated above [Almond2015], the learner's state of knowledge after the instruction is dependently on the instruction itself, as well as learner's readiness for it. Our evidence of learning only tells us about measurements implying extent of learning about prior learning instructions. Watson Education cannot therefore suggest inferences about future performance.

Finally, the fourth question — treading tenuously close to that avoided above — suggests extrapolation of future trends. That past performance often indicates future performance is a common experience for Education teachers. Watson-Education users can observe trends and form *opinions* about their implications for the individual students' future learning trajectories. *Such opinions are however altogether unsupported by Watson-Education's AI-based DCM*.

This particularly applies to students exhibiting mixtures of strength and weakness in their trajectories through the learning-standard progressions. Section 2 explicitly addresses phenomenological obstacles to employing AI-algorithmic logic in projecting, forecasting, or predicting learners' future states of knowledge with degrees of certainty that might justify the required effort. Watson-Education's Mastery feature vigilantly avoids asserting bases for belief about students' knowledge and readiness states regarding which the underlying degree of uncertainty cannot be clearly explained.

1.2.3 Evidence-centered design (ECD).

Evidence-Centered Design (ECD) originated in the educational-measurements community. Designing educational assessments that elicit precise, high-confidence evidence about the test subject's state of knowledge with respect to specific competencies [Mislevy2003, Almond2015, Haertel2016]. Educational assessments represents its original intended application.

Figure 3 depicts how IBM adapted ECD to competency-based personalized learning. Watson Education extends Mislevy's original ECD-model view [Mislevy2003, Figure 1] by adding a learner model. Incidentally, Haertel, *et al*, incorporate a Learner model in their ECD point of view.

IBM renders its ECD point of view in Figure 3 employing notation conforming to a Unified Modeling Language (UML) Class Diagram [Podeswa2005]. The major ECD–model components are designated using the classifier notation, such as ⟨⟨Learner Model⟩⟩ for the learner model. The classifier associations — aggregation associations and association classes — depict IBM's distinctive point of view on the workflow for competency-based personalized learning. The association directions represent IBM's view of competency-based learning's operational flow.

Competency-based personalized learning starts with the learner. We capture the essential elements about the learner in the ⟨⟨Learner Model⟩⟩. This is congruent with the learner-centered perspective of Figure 1.1. In fact, the Learner-Model attributes in Figure 3 coincide actually with the concentric rings of Figure 1.1. The "Outcomes Sought" and the "Knowledge, Readiness States" pertain specifically to competency-based learning.

Although Mislevy originally thought of the ⟨⟨Assembly Model⟩⟩ as the assessment-administration system, it pertains in IBM's point if view to the to the curriculum framework through with competency-based instruction is delivered. The ⟨⟨Assembly Model⟩⟩ includes both the curriculum trajectory, characterized by "Course Progression" as well as individual academic courses represented by "Course Administration". Aspirationally, Watson-Education supports differentiated ⟨⟨Assembly Model⟩⟩ realizations. These are distinguished by the college and career objectives pursued by the individual learner. Watson-Education supports only high-school graduation at the time of this writing.

The ⟨⟨Proficiency Model⟩⟩ specifies competencies the learner ought to have mastered upon completion of instruction. Watson Education's ⟨⟨Proficiency Model⟩⟩ consists at the time of this writing of state learning standards. It leverages the

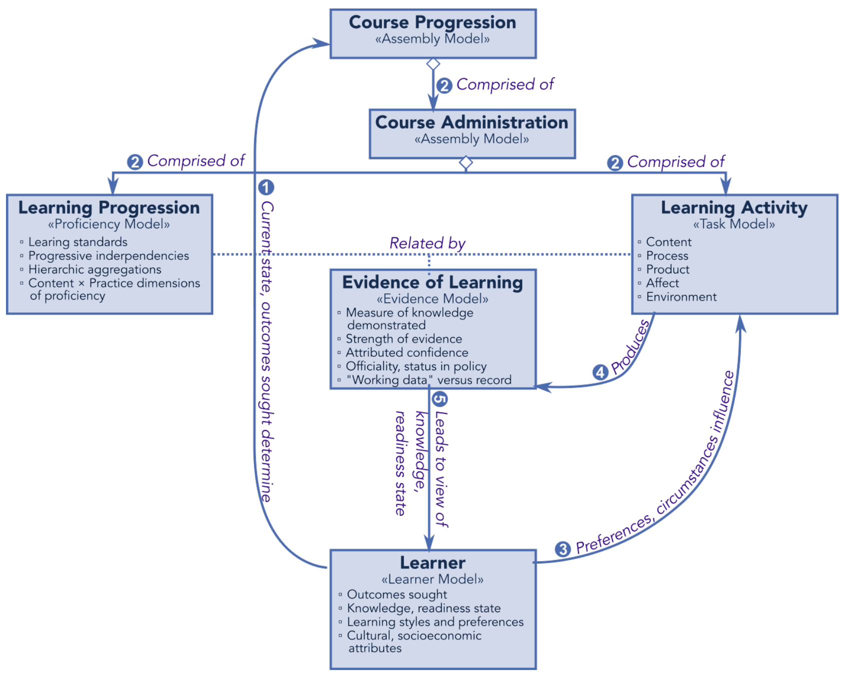

Figure 3: Mislevy's Evidence-Centered Design (ECD) framework exerts significant influence on Watson-Education's concept of operations for competency-based personalized learning. (After: [Mislevy2003, Figure 1]).

learning-standards "infrastructure" states deployed in order comply with federal accountability requirements under the *Elementary and Secondary Education Act* (ESEA) [ESEA2001]. Watson-Education's ⟨⟨Proficiency Model⟩⟩ also includes learning-standard vertical articulations.

Watson-Education's Mastery features answers questions like "What is the the learner's state of knowledge?" Or, "Is the learner ready for the instructional content planned next?" The ⟨⟨Proficiency Model⟩⟩ adds the "With respect to what?" about each of the questions. Learning-standard vertical articulations contain the relationships Watson Education uses to infer the answers.

The ⟨⟨Proficiency Model⟩⟩ provides the basis for design of the ⟨⟨Task Model⟩⟩. The ⟨⟨Task Model⟩⟩ captures learners' activities producing evidence of learning. Almond and Mislevy [Almond2015] describe a "design-forward" framework for deriving tasks from proficiency specifications. Wiggin's and McTighe's conceptually congruent "Understanding by Design" framework [Wiggins2005] uses the term "backward design". Learning activities are carefully designed to elicit evidence of the learner's extent of mastery of learning targets from the ⟨⟨Proficiency Model⟩⟩ .

The attributes in Watson Education's ⟨⟨Task Model⟩⟩ are based on Carol Ann Tomlinson's differentiated-instruction framework [Tomlinson2014]. This framework when carefully followed defines learning activities that are aligned to learning objectives from the ⟨⟨Proficiency Model⟩⟩. IBM conscientiously selected this generalized framework for eliciting evidence of learning from learners. Incidentally, both Tomlinson and Almond refer to actual evidence-bearing outputs from students as "Work Products."

The ⟨⟨Evidence Model⟩⟩ model associates the ⟨⟨Proficiency Model⟩⟩ with the ⟨⟨Task Model⟩⟩. It includes Evidence of Learning produced by the ⟨⟨Task Model⟩⟩. It also contains the logic by which Evidence of Learning is translated into a belief about individual learners' state of knowledge. Individual learners' states of knowledge with respect to the ⟨⟨Proficiency Model⟩⟩ are contained in the ⟨⟨Learner Model⟩⟩. The latter, described in [Mislevy2003, Almond2015, Haertel2016], is not explicitly depicted in Figure 3.

1.2.4 Learner-centered curriculum (*aka* "curriculum alignment").

Curriculum alignment represents the second operational framework from which Watson-Education concept of operations is derived. Curriculum alignment is based from *Opportunity to Learn* (OTL) theory. To state things very simply, OTL asserts that learners can only acquire knowledge from educational instruction that is actually contained in the instruction itself [Elliott2016], [Walkowiak2017].

University of Pennsylvania education researcher Andrew Porter developed a framework to quantify the extent to which curriculum provides learners with an opportunity to learn [Porter2001]. Figure 4 depicts IBM's point of view on the application of Porter's framework to competency-based personalized learning. It contains the four facets of an aligned curriculum described by Porter. It elaborates on them however in terms of actual entities IBM finds in curriculum, instruction, and assessments (CIA) frameworks employed by Watson-Education school-district clients. It also depicts the Watson-Education information flow in a manner similar to that for ECD in Figure 3.

ECD and curriculum alignment are congruent and complementary views of curriculum, instruction, and assessment (CIA) in education. ECD emphasizes the assessment aspect. Curriculum alignment "wraps" the curriculum and instruction around things.

Porter's curriculum-alignment contains describes four facets of an aligned curriculum. The *Intended Curriculum* refers to such policy tools as curriculum standards, frameworks, or guidelines that outline the curriculum teachers are expected to deliver. In addition to the learning standards themselves, IBM looks for the scope and structure of academic courses. The scope is specified in terms of learning standards addressed by the course. Its structure includes instructional units (or modules). We also employ a vertical articulation (sometimes also referred to as a coherence map) of learning standards within and across courses. The Intend Curriculum corresponds to the ECD ⟨⟨Assembly Model⟩⟩ and ⟨⟨Proficiency Model⟩⟩ .

The *Enacted Curriculum* refers to the actual curricular content that students engage in the Education. The intended, assessed, and learned curricula are important components of the educational delivery system, but most learning is expected to occur within the enacted curriculum. The enacted curriculum is comprised of instructional content, including materials curated in the Watson Education Library feature. It also includes pedagogical practices individual teachers apply in distinct courses and sessions. The Enacted Curriculum contains the ECD ⟨⟨Task Model⟩⟩ .

The *Assessed Curriculum* measures the extent to which an individual student masters each distinct learning distinct learning objective. Watson Education focuses here on learning standards. Watson Education's feature design for this was influenced by concepts from education-instruction authorities Carol Ann Tomlinson, and from Jay Wiggins and Grant McTighe. The Mastery diagnostic feature considers all available criterion-referenced (standards-aligned) evidence of learning in providing teachers with a view of learners' strengths and weaknesses.

The *Learned Curriculum* is that portion of the intended curriculum that the learner has actually mastered. Figure 5 elaborates on Figure 4 as to how the Assessed Curriculum leads to inferences about the Learned Curriculum. Watson Education's feature represents an attempt to synthesize a view of the learned curriculum for each student given all-available evidence of learning. This diagnostic seeks to represent each learner's knowledge state with respect to each learning objective. The ECD ⟨⟨Learner Model⟩⟩ contains the Learned

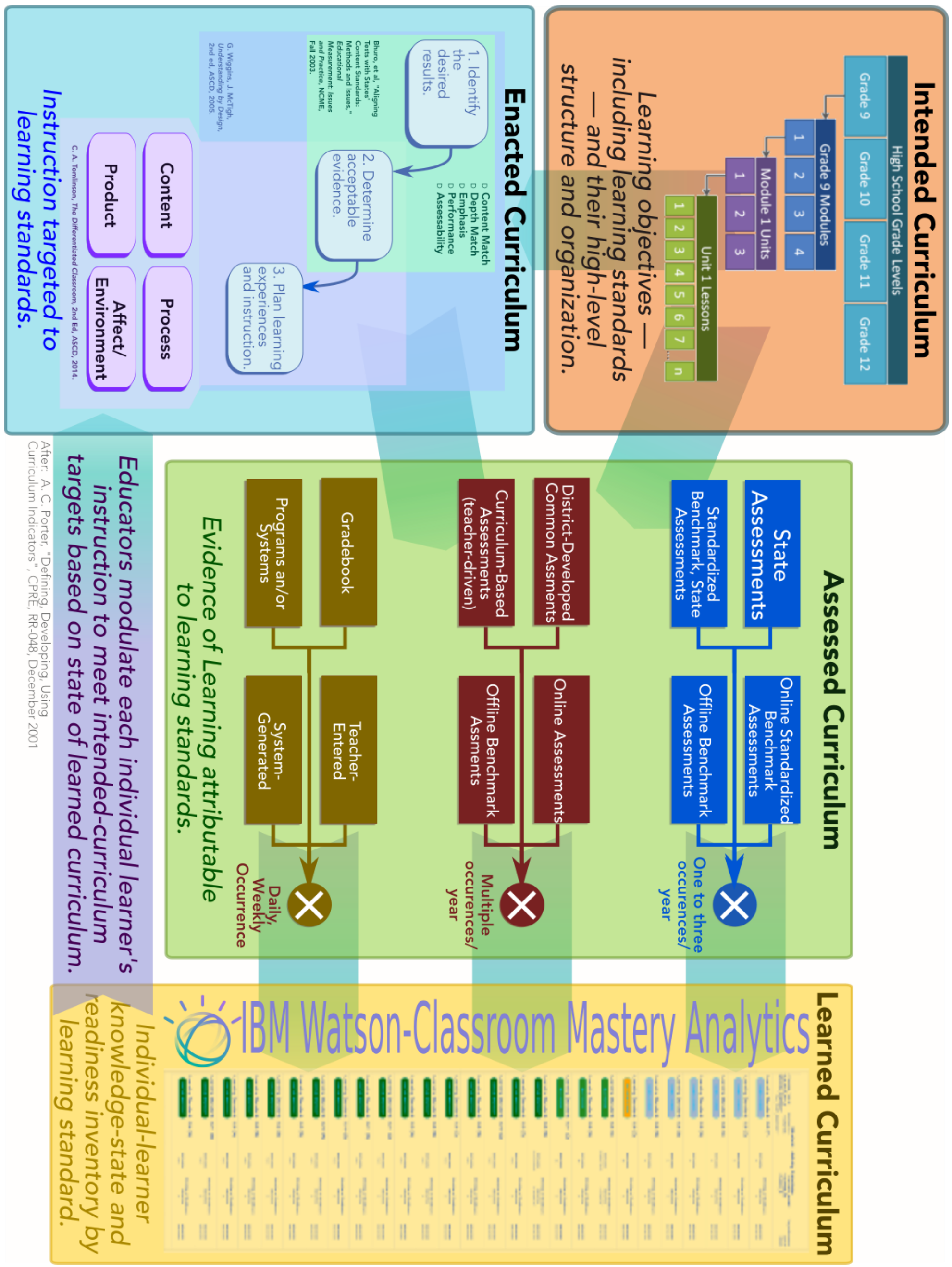

Figure 4: Competency-based personalized learning is best-accomplished through a well-aligned curriculum. (After: [Porter2001, Figure 3].)

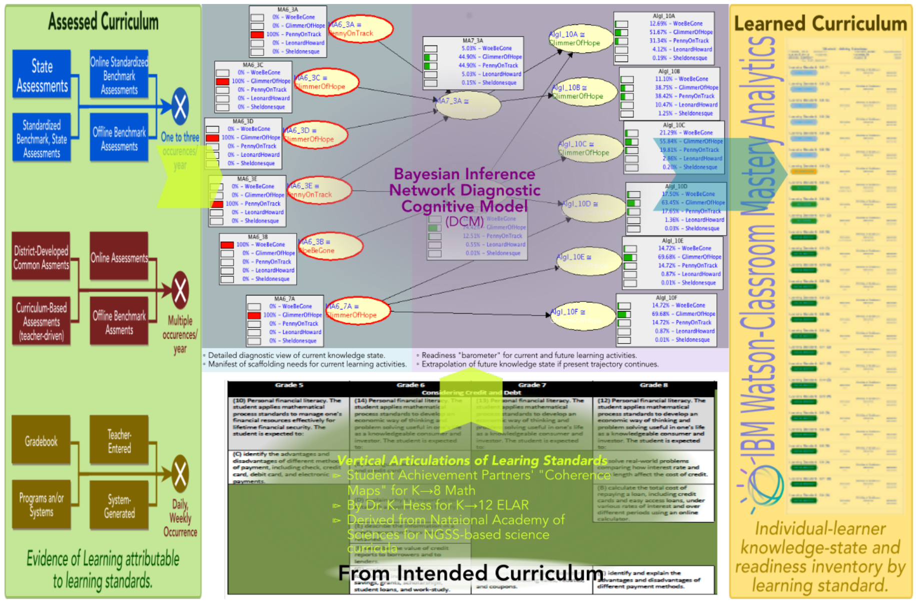

Figure 5: The relationship between the assessed and learned curricula links curriculum alignment Figure 4 to ECD Figure 3.

Curriculum. Of particular note, Learning-standard vertical articulations — described in section 1.2.1 — capture education researchers' views of how learning progresses.

The heart of Watson Education Mastery feature lies with the association between the Assessed Curriculum and the Learned Curriculum. Figure 5 illustrates the logic by which evidence of knowledge from the Assessed Curriculum is translated into estimates of the Learned Curriculum. Figure 5 explicitly shows the Bayesian Network, mechanism Watson Education employs to link the Assessed Curriculum to the Learning Curriculum. Entire sections of this document subsequently describe the algorithmic details.

IBM observes here that the education-assessments community exercises great care to distinguish between test scores, produced in the Assessed Curriculum, and what is actually learned. The former are imperfect manifestations of the latter. A rich theory of cognitive diagnostics guides the design of "instruments" (assessments) and interpretation of their results. Included among the published titles influencing IBM's Mastery solution are [Almond2015], [Loehlin2004], [Rupp2010], [Leighton2007], [Reckase2009], and [DeMars2010]. This document elaborates at length in Section 2 on education-measurements phenomenology.

1.3 Watson-Education concept of operations for Mastery and competency-based personalized learning.

Figure 6 puts all of this together into a moderately-detailed view of Watson Education's concept of operations — or theory of action. Watson Education aspires to help educators adaptively and agilely manage complexity and variation in their student populations. We seek to put at teachers' fingerprints a clear picture of their students' state of progress through the curriculum. Considerably more variation exists than traditional, time-based instructional frameworks [Bramante2012] easily deal with.

Watson-Education is designed to help Education teachers use a rapid decision-feedback loop to quickly and efficiently manage the variability among students. Such frameworks — generally referred to as "learning loops" — have been widely employed by in complex business and military operations. Descriptions of their use appear in [Haeckel1993], [Partnoy2012], [Bonchek2013], and [Feloni2017].

Figure 7 introduces learning loops at a cursory level. Two versions are depicted. The original concept — shown on the left — finds its roots in military arial combat. Decision-feedback cycles follow an Observe-Orient-Decide-Act (OODA) trajectory. The right-hand rendering — Sense-Interpret-Decide-Act — is a demilitarized version for business purposes.

IBM's renderings of ECD in Figure 3 and curriculum alignment in 4 employ this type of decision-feedback logic. Figure 7 contains abstractions of the Watson-Education concept of operations. Teachers employ use Watson-Education to engage in decision-feedback thinking when they recognize an individual student's degree of readiness and modulate their planned instruction accordingly.

1.3.1 *Observing* or *Sensing* stage of learning loop.

Observing or *Sensing* when occurs, for example, when a teacher recognizes that Watson Education shows a student is not ready for the content about to be presented. Education teachers see at a glance in Watson Education's Mastery presentation the readiness of each student to engage the instructional content about to be presented. They enjoy detailed and comprehensive situational awareness.

Figure 8 depicts an example. Consider a teacher teacher who is about to present content related to learning standard "M07.C-G.2.2.1". The "Class Breakdown" pane indicates Watson's estimate of each student's readiness to engage the content. The "Class Breakdown" view shows the state of readiness for each learner in enrolled in the class. That a significant swath of students might not be prepared is the essential information provided by Watson-Education Mastery.

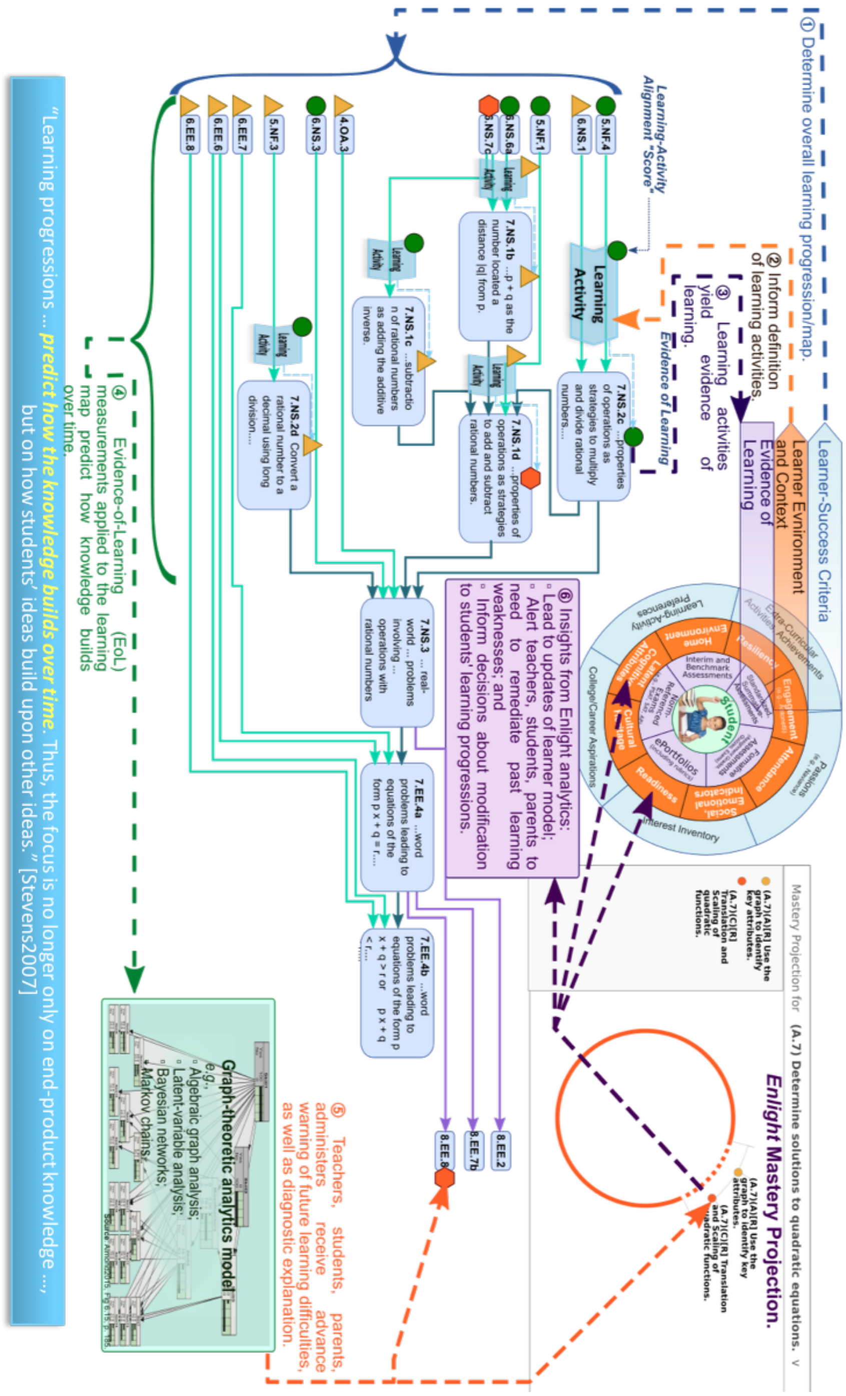

Figure 6: The Concept of Operations for IBM Watson Education helps teachers guide learners through their curriculum through a framework based on decision-feedback cycles.

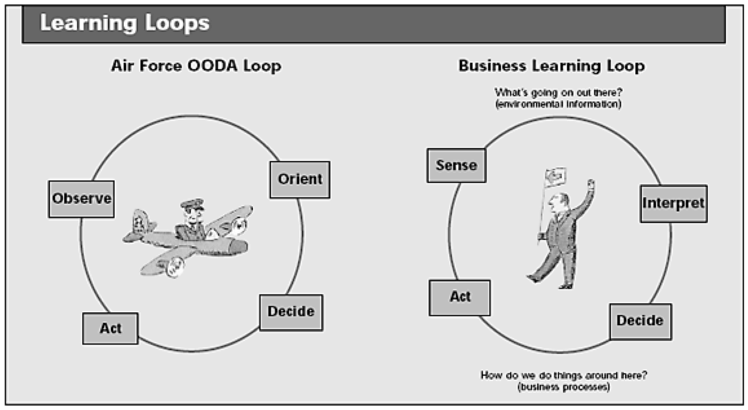

Figure 7: A "*Learning loop*" (From: [Haeckel1993]) is a simple, general form of a decision-feedback cycle that is commonly applied to operational decision-making in dynamic, complex environments. Managing the differentiation of instruction or personalization of learning for a Education of learners of varying degrees of readiness has exactly the level of dynamic complexity typical for learning-loop applications.

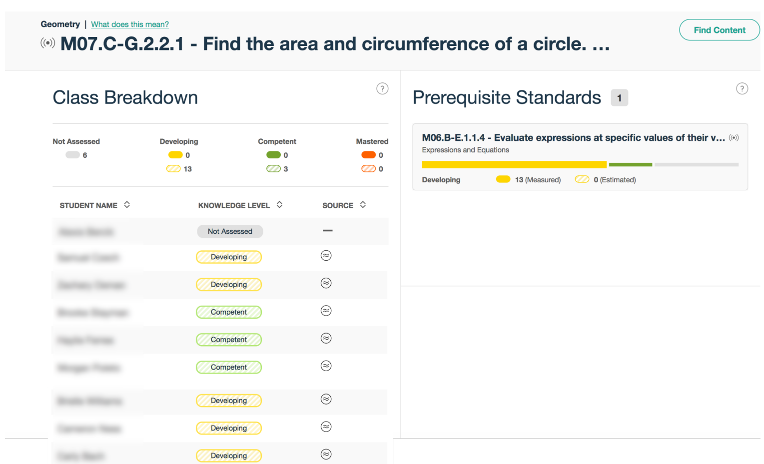

Figure 8: The Watson-Education presentation for Mastery gives teachers a view at a glance of each student's readiness to engage instructional content about to be presented.

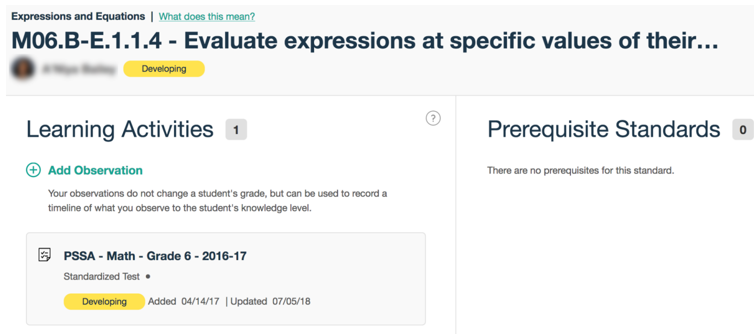

Figure 9: Watson-Education Mastery allows teachers to "drill back" along the learning-standard progressions to understand the evidence of learning on which Watson's view of a student's state of knowledge is based. Illustrate here is the "drilling back" into an individual *learner's evidence of learning* on which the *class-readiness view* of Figure 8 is based.

1.3.2 *Orienting* or *Interpreting* stage of learning loop.

The *Orient* or *Interpret* stage of the learning loop next occurs. Our teacher here relates observations about students' readiness to the operational context. In this case, the instruction about to be presented constitutes the context. Specifically, Watson Education attempts to answer the question, "Why are some of my students unready?"

Watson answers this question in terms of terms of proficiency. Specifically, learning standard "M07.C-G.2.2.1" builds on knowledge from prerequisite learning standard "M06.B-E.1.1.4". This information appears in the "Prerequisite Standards" pane of the Watson-Education display. The prerequisites are specified by the learning-standard progressions from section 1.2.1. Only one prerequisite is specified for this illustration. Many learning standards however have multiple prerequisites.

To answer the essential Orient/Interpret question about "Why", our teacher clicks the card corresponding to prerequisite learning standard "M06.B-E.1.1.4". This provides a view for "M06.B-E.1.1.4" exactly like that for "M07.C-G.2.2.1" shown in Figure 8. The teacher sees the state of knowledge for each student with respect to "M06.B-E.1.1.4". This amounts to orienting each students' state of knowledge — their Learned Curriculum — with respect to the Intended Curriculum.

Teachers also "drill into" individual students' states of knowledge. Figure 9 illustrates. The teacher sees the actual learning activities on which Watson's diagnostic was based. The figure illustrates a scenario in which only one evidentiary measurement (learning activity) is available. If multiple measurements are present, Watson combines their scoring by a category-weighted average.

1.3.3 *Decide* and *Act* stages of learning loop.

In terms of curriculum alignment (Figure 4), Education teachers' primarily influence the Enacted Curriculum. The *Decide* and *Act* stages of the learning loop pertain therefore to teachers' execution of the Enacted Curriculum. The Intended Curriculum is often provided Education teachers by higher authority, such as states or school districts. These policies often specify learning-standard scope and can include rough allocations of Education time for each learning standard.

Teachers' decision prerogatives include minor adjustments to scope, adjustments to Education-time allocations, pedagogical approach, introduction of supplementary instructional materials, evidentiary interpretation of evidence-of-learning work products (grading), and formative-assessment approaches. Watson Education supports decisions and actions within this scope.

Watson Education's Library Services feature is particularly useful. Figure 10 the Library Services feature. A teacher queries Watson Education to get a list of instructional materials for a specified learning standard. This has multiple applications.

Supplementary materials can be pulled into instructional plans for selected students. These materials can provide scaffolding to review or reinforce prerequisite learning standards regarding which not-completely-prepared students are weak. Additional, more-challenging materials can alternatively be offered to challenge students for whom the baseline instruction might not be sufficiently strenuous.

1.3.4 Learning-loop feedback: *Observing* outcomes from *Acting*.

The feedback stage of the learning loop links Observe/Sense activities back to Act. Teachers' delivery of instructional materials to students produce learning-activity work products [Wiggins2005, Tomlinson2014]. These work products represent evidence of learning. This evidence can be either the evaluated work product itself, or a supplementary formative or summative assessment.

Watson Education's Mastery algorithm updates its view of each student's state of knowledge given this new evidence. This provides teachers with feedback as to how effective their enactment of the curriculum has been. They can in response make decisions accordingly to modulate subsequent instruction.

For example, academic courses often include "spiraling", by which specific learning standards are revisited multiple times. Students' who appear to master a learning standard on the first attempt might require less exposure on subsequent visits. This opens up flexibility to allocate Education time to other topics. Alternatively, additional scaffolding or other emphasis can be built into subsequent visits to a learning standard by students who didn't quite get it down on the first pass.

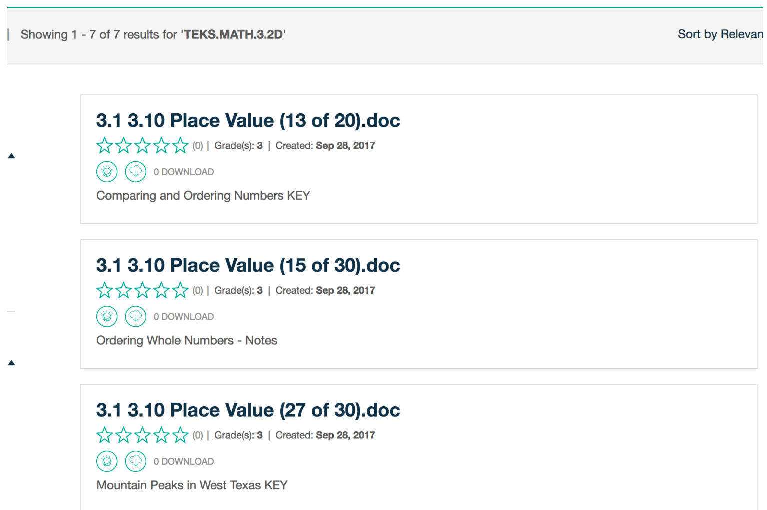

Figure 10: Watson Education's Library Service curates learning-standards-aligned instructional materials by which teachers can modulate the scope instructional approach of the Enacted Curriculum.

2 Education-measurements phenomenology and the data they produce.

2.1 Phenomology model for learning.

2.1.1 Phenomenology of learning.

Figure 11 depicts our phenomenology model for learning. It shows a contrived, illustrative learning-progression scenario. Educational-curriculum specialists tell us that learning standard 2.A builds upon knowledge specified by three prerequisite learning standards, 1.A, 1.B, and 1.C. The directed graph in the yellow-shaded domain depicts this. This representation is equivalent to that in Figure 2.

It depicts Stevens' basic idea from section 1.2.1: "Learning progressions describe how students gain more expertise within a discipline over a period of time. They represent not only how knowledge and understanding develops, but also predict how the knowledge builds over time. Thus, the focus is not limited on the end-product knowledge as characterized by summative assessment, but on how students' ideas build upon other ideas" [Stevens2007]. It also reflects Almond's hypothesis that, "The student's probability of knowing the skill at the end of the course will depend both on the student's skill level at the beginning of the course, and what kind of instruction the student receives" [Almond2015, p. 98].

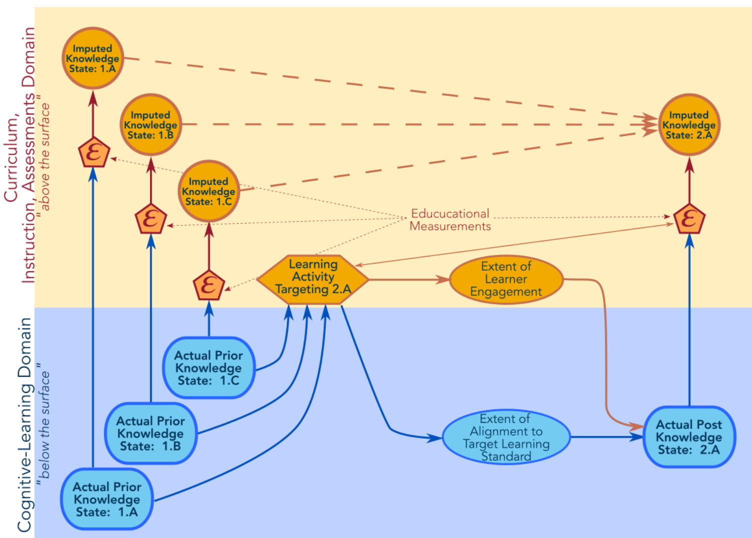

Figure 11: Ascertaining the "Learned Curriculum" (Figures 4 and 5) for an individual learner involves using "*above-the-surface*" educational measurements $\mathcal{E}$ to make inferences about not-directly-observable, "*below-the-surface*" states of knowledge.

Figure 11 also conveys our foundational view of Mastery as a Diagnostic Cognitive Model (DCM) [Leighton2007, Rupp2010]. It also shows the progressive nature of learning, intruduced in Section 1.2.1.

We partition our view of the phenomonology into two domains. The "*above-the-surface*" domain contains those things that educators can observe directly. This domain spans the realm of Curriculum, Instruction, and Assessments (CIA). The observables domain is shaded in yellow.

The latent — or "*below-the-surface*" — domain contains unobservable aspects of our phenomonology. Shaded with a blue background, this contains the phenomenology associated with cognitive learning. Whatever educators observe above the surface are manifestations of students' state of knowledge. Knowledge states reside below the surface.

In terms of our curriculum-alignment framework in Figures 4 and 5, the above-the-surface region corresponds to the Intended, Enacted, and Assessed facets of the curriculum. The learning standards themselves — 1.A, 1.B, 1.C, and 2.A — are specified by the Intended Curriculum. The Enacted Curriculum contains the learning activity itself. The evidence of learning for the prerequisite standards comes from the Assessed Curriculum. Watson-Education Mastery's DCM estimates the Learned Curriculum, which appears in the below-the-surface domain of Figure 11.

Our scenario focuses on a learning activity targeting learning standard 2.A.

Before starting this learning activity we have prior evidence of learning about learners' knowledge states for prerequisite learning standards 1.A, 1.B, and 1.C. Our belief about each learner's knowledge states for these prerequisites determines our belief about her readiness to engage the learning activity.

We have no evidence about learners' knowledge of learning standard 2.A at this point. Watson Education's DCM estimates each learner's readiness for the learning activities addressing 2.A. In terms of Figure 2, our measurements about prerequisite learning standards 1.A, 1.B, and 1.C are represented in the blue-shaded region. Our readiness estimate for 2.A is in the purple-shaded region.

We use the symbol ε to depict an educational measurement from the assessed curriculum. The use of the Greek letter ε here is a double entendre. First, ε represents evidence applied to our Bayesian-Inference Network (BIN) DCM. We introduced this model in Figure 5. Subsequent sections go into extensive detail. The key point here is that evidence ε provides an indirect indication of a latent — unobservable — state of knowledge. Section 2.2 elaborates at length.

Second, ε commonly denotes error in statistical models. Denoting our evidence this way reminds us that our measurements are accompanied by difficult-to-characterize uncertainty and imprecision. Our evidentiary measurements — assessment results or other evidence of learning — are never true, exact indications of the learner's state of knowledge.

Now, Almond's hypothesis [Almond2015, p. 98] asserts that the learner's actual post-instructional knowledge state depends on three things. The *actual, prior knowledge states for prerequisite learning standards* represents the first. Our assessed (or observed) knowledge states are based on indirect evidence ε. The uncertainty in ε introduces the potential for ambiguity about the learner's actual readiness.

Secondly, Almond addresses the quality of instruction. The instruction itself must be aligned to the targeted topic. Instructional alignment is nontrivial. [Bhola2003], *et al*, describe specific frameworks. There can be three to five relevant degrees of alignment. Pedagogy — the art of effective presentation — is also important.

To Almond's hypothesis we add the importance of learner engagement. Antonetti and Garver [Antonetti2005] developed a comprehensive framework for characterizing learner engagement. Figure 12 presents the "Antonetti Cube". Learner engagement is obviously a complex discipline in and of itself. Further elaboration lies beyond the scope of this document.

In view of Figure 12, clarifying what learner engagement is not becomes important. Some education-technology vendors refers to learner engagement in strikingly-different terms. Digital-instruction platforms seek to measure "engagement" using online-platform session statistics.

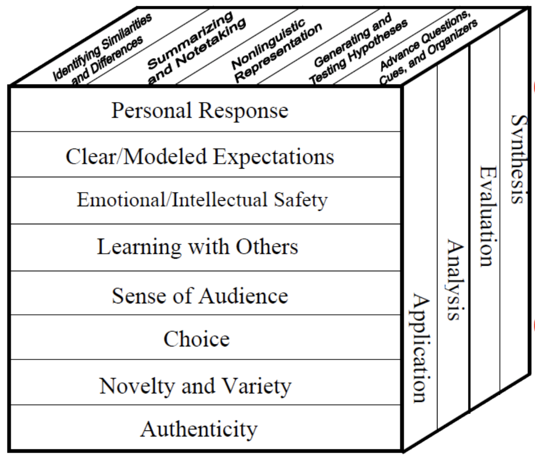

Figure 12: Learners' states of knowledge upon conclusion of instruction depends on the extent of their engagement with the instructional content. (From [Antonetti2005].)

Other research [Cao2012] explored monitoring eye movements to make inferences about learning styles. Other research (e.g., [Bailey2018]) however suggests alternative explanations for apparent divided attention than loss of opportunity to learn based on weak engagement. Observables used by Education-Technology vendors may have alternative explanations to hypotheses about efficiency of instruction.

What is the upshot of all of this? Cognitive learning is characterized by a deep phenomenology. This section makes that case. We further elaborate in in the next, with a view on the "architecture" of knowledge.

Watson-Education Mastery only has access to two sources of information about individual learners' states of knowledge. Mastery uses educational measurements. We suggest here — and make the case more-deliberately in section 2.2 — that education measurements are at best silhouettes of the underlying congnitive states of knowledge. We also have — external to our data themselves — curriculum experts's beliefs as to the the progressions by which pieces of knowledge build upon one another.

Figure 13 boils it down to the phenomenology to which Watson-Education Mastery has access. We are limited to educational measurements that

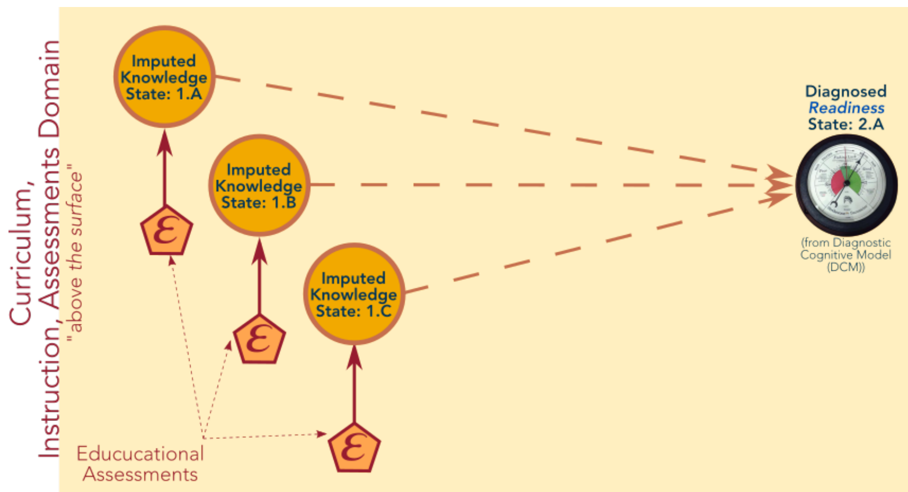

Figure 13: Watson-Education Mastery's DCM translates *imputed knowledge states* into *diagnosed readiness states* using a Bayesian Inference Network ([Rupp2010, section 5.4]) method.

2.1.2 "Architecture" of knowledge.

2.2 Characteristics of education-measurements data.

References

[NCSL2017] "Competency-based education", National Council of State Legislators, September 8, 2017, http://www.ncsl.org/research/education/competency.aspx.

[Bramante2012] F. Bramante and R. Colby, *Off the Clock: Moving Education from TIME to COMPETENCY*, Thousand Oaks, CA: Corwin, 2012, https://amzn.to/2A06iuK.

[USDoEd] "Competency-Based Learning or Personalized Learning", U.S. Department of Education, https://www.ed.gov/oii-news/competency-based-learning-or-personalized-learning.

[Tomlinson2014] C. A. Tomlinson, *The Differentiated Education*, second edition, Alexandria, VA: ASCD, 2014, http://bit.ly/2mTVNjr.

[Wiggins2005] G. Wiggins, J. McTighe, *Understanding by Design*, second edition, Alexandria, VA: ASCD, 2005, http://bit.ly/2uZvhIK.

[Rickabaugh2016] J. Rickabaugh, *Tapping the Power of Personalized Learning*, Alexandria, VA: ASCD, 2016, http://bit.ly/2A3Ums5.

[Marzano2008] R. J. Marzano and M. W. Haystead, *Making Standards Useful in the Education,* Alexandria, VA: ASCD, 2008, http://bit.ly/2A3gUZN.

[Drake2004] S. M. Drake and R. C. Burns, *Meeting Standards through Integrated Curriculum,* Alexandria, VA: ASCD, 2004, http://bit.ly/2AoO8D3.

[Flemming2001] N. Flemming, *Teaching and Learning Styles*, Christchurch, NZ, 2001.

[Khazan2018] O. Khazan, "The Myth of 'Learning Styles'", *The Atlantic*, April 11, 2018, https://www.theatlantic.com/science/archive/2018/04/the-myth-of-learning-styles/557687/.

[Murray1976] Murray, N. J. M. (1976). COMPETENCY-BASED LEARNING PACKAGES - A CASE STUDY. *Training and Development Journal*, 30(9), 3.

[SRA] SRA Mathematics Laboratory, McGraw Hill, https://amzn.to/2uWXEYo.

[Farrington2012] C. A. Farrington, *et al*, *Teaching Adolescents To Become Learners: The Role of Noncognitive Factors in Shaping School Performance: A Critical Literature Review*, University of Chicago Consortium on Chicago School Research, June 2012, http://ibm.biz/UCCCSR-Non-Cognitive.

[Edelmen2015] D. C. Edelmen and M. Singer, "Competing on customer journeys," *Harvard Business Review*, November 2015, https://hbr.org/2015/11/competing-on-customer-journeys.

[Christensen2016] C. Christensen, *et al*, "Know your customers' 'jobs to be done'", *Harvard Business Review*, September 2016, https://hbr.org/2016/09/know-your-customers-jobs-to-be-done.

[Manning2012] H. Manning, K. Bodine, and J. Bernoff, *Outside-In: The Power of Putting Customers at the Center of Your Business*, Amazon Publishing, August 28, 2012, https://amzn.to/2A84iRm.

[Kane2015] G. C. Kane, et al, "Strategy, not technology, drives digital transformation", *MIT Sloan Management Review*, July 14, 2012, https://sloanreview.mit.edu/projects/strategy-drives-digital-transformation/.

[LaValle2011] S. Lavalle, et al, "Big data, analytics, and the path from insights to value," *MIT Sloan Management Review*, Winter 2011, https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/.

[Satell2018] G. Satell, "How to make an AI project more likely to succeed", *Harvard Business Review*, July 19, 2018, https://hbr.org/2018/07/how-to-make-an-ai-project-more-likely-to-succeed.

[SAMIAM] Sensitivity, Analysis, Modeling, Inference, and More (SAMIAM), University of California at Los Angeles, 2010, http://reasoning.cs.ucla.edu/samiam/.

[CCSS-Math] *Common Core State Standards: Mathematics*, Common Core State Standards Initiative, Council of Chief State School Officers (CCSSO) and National Governors Association, http://www.corestandards.org/Math/.

[Almond2015] R. Almond, *et al*, *Bayesian Networks in Educational Assessments*, New York: Springer, 2015, https://www.springer.com/us/book/9781493921249.

[Stevens2007] Stevens, S., *et al*, (2002), "Using learning progressions to inform curriculum, instruction and assessment design," Center for Highly Interactive Educations, Curricula & Computing in Education, School of Education/College of Engineering, University of Michigan, Retrieved on July 24, 2018 from http://ibm.biz/CHICCCE-Progressions.

[ESEA2001] "No Child Left Behind Act of 2001", *Elementary and Secondary Education Act* (ESEA), Public Law 107-110, https://www2.ed.gov/policy/elsec/leg/esea02/index.html.

[Askey] R. Askey (reviewer), et al, "Progressions documents for the Common Core math standards", Instititue of Mathematics and Education, University of Arizona, Funded by Brookhill Foundation, http://ime.math.arizona.edu/progressions/.

[Zimba2012] J. Zimba, "A graph of the content standards", Student Achievement Partners, June 7, 2012, http://ibm.biz/Zimba-Wiring-Diagram.

[SAP] Coherence Map, Student Achievement Partners, https://achievethecore.org/page/1118/coherence-map.

[DLM] Dynamic Learning Maps, University of Kansas, https://dynamiclearningmaps.org/.

[NGSSApdxE] “Progressions within the Next-Generation Science Standards”, Next-Generation Science Standards Appendix E, http://ibm.biz/NGSS-ApdxE.

[tenBerge1999] T. ten Berge and R. van Hezewiyk, “Procedural and declarative knowledge”, *Theory & Psychology,* Sage, October 1999, http://ibm.biz/TheoryNPsych-KnowledgeTypes.

[Yilmaz2012] İ Yilmaz and N. Yalçin, “The relationship of procedural and declarative knowledge of science-teacher candidates in Newon’s laws of motion to understanding,” *American International Journal of Contemporary Research*, March 2012, http://ibm.biz/AIJCR-Sci-Know-Types.

[NYNGMath] New York State Next-Generation Mathematics Learning Standards, New York State Department of Education, http://ibm.biz/NYNGLS-Math.

[TXGateway] “Vertical Alignment Charts for Revised Mathematics TEKS”, Texas Gateway, http://ibm.biz/TEKS-Math-Vert-Algmt.

[TEA2017] Expert Reviewer Recommendations, English, Language Arts, and Reading (ELAR) vertical alignments, December 2016, https://tea.texas.gov/Academics/English_TEKS_Review/.

[Hess2011] K. Hess and J. Kearns, “Learning Progressions Frameworks Designed for Use with The Common Core State Standards in English Language Arts & Literacy K-12”, National Center for the Improvement of Educational Assessment (NCIEA), December 2011, http://ibm.biz/NCIEA-CSSS-ELA-Progressions.

[Mislevy2003] R. J. Mislevy, et al, “A brief introduction to evidence-centered design,” Research Report RR-03-16, Educational Testing Service (ETS), July 2003, https://www.ets.org/Media/Research/pdf/RR-03-16.pdf.

[Haertel2016] G. Haertel, *et al*, “General introduction to evidence-centered design”, *Meeting the Challenges to Measurement in an Era of Accountability*, New York: Routledge, for the National Council on Measurement in Education, 2016, https://amzn.to/2OhMMx7.

[Podeswa2005] H. Podeswa, *UML for the IT Business Analyst*, Boston, MA: Thompson, 2005, https://amzn.to/2A73vQC.

[Elliott2016] S. N. Elliott and B. J. Bartlett, "Opportunity to Learn", *Oxford Handbooks Online*, May 2016, http://ibm.biz/Oxford-Hbk-OTL.

[Walkowiak2017] T. A. Walkowiak, et al, "A reconceptualization framwork for 'Opeortunity to Learn in school mathematics," Jornal of Mathematics Education at Teachers College, Teachers College, Columbia University, Spring 2017, http://ibm.biz/JMETC-OTL.

[Porter2001] A. C. Porter and J. L. Smithson, "Defining, developing, and using curriculum indicators," Report RR-048, Consortium for Policy Research in Education, University of Pennsylvania, December 2001, http://ibm.biz/PorterCurriculumAlignment.

[Loehlin2004] J. C. Loehlin, *Latent Variable Models*, fourth edition, Mahwah, NJ: Lawrence Erlbaum Associates, 2004, https://amzn.to/2AkjOJQ.

[Rupp2010] A. Rupp, *et al*, *Diagnostic Measurement: Theory, Methods, and Applications*, New York: Guilford, 2010, https://amzn.to/2AnMHVy.

[Leighton2007] J. P. Leighton and M. J. Gierl, *Cognitive Diagnostic Assessment for Education*, London: Cambridge Univeristy Press, 2007, https://amzn.to/2Apb0Cs.

[Reckase2009] M. D. Reckase, *Multidimensional Item Response Theory*, New York: Springer, 2009, https://www.springer.com/us/book/9780387899756.

[DeMars2010] C. Demars, *Item Response Theory*, London: Oxford University Press, 2010, https://amzn.to/2AoFYe7.

[Haeckel1993] S. H. Haeckel and R. Nolan, "Managing by wire", *Harvard Business Review*, September-October 1993, https://hbr.org/1993/09/managing-by-wire.

[Partnoy2012] F. Partnoy, "Act fast, but not necessarily first", *Harvard Business Review*, July 13, 2012, https://hbr.org/2012/07/act-fast-not-first.

[Bonchek2013] M. Bonchek and C. Fussel, "Decision making, Top Gun style", *Harvard Business Review*, September 12, 2013, https://hbr.org/2013/09/decision-making-top-gun-style.

[Feloni2017] R. Feloni and A. Pelisson, "A retired Marine and elite fighter pilot breaks down the OODA Loop, the military decision-making process that guides 'every single thing' in life'", *Business Insider*, August 13, 2017, https://read.bi/2mJSlHR.

[Bhola2003] D. Bhola, *et al*, "Aligning tests with states' content standards: Methods and issues," *Education Measurement: Issues and Practices*, National Council on Measurements in Education, Fall 2003, http://ibm.biz/NCME-ExamAlgmt.

[Antonetti2005] J. V. Antonetti and J. R. Garver, *17,000 Classroom Visits Can't Be Wrong*, Springfield, VA: Association for Supervision and Curriculum Development, 2015, http://bit.ly/2olb3Gw.

[Cao2012] J. Cao and A. Nishihara, "Understanding learning style by eye tracking in slide-video learning, *Journal of Educational Multimedia and Hypermedia,* November 2012, http://ibm.biz/JEMMH-EyeMovement.

[Bailey2018] C. Bailey, "Distracted? Work Harder!", *New York Times*, August 25, 2018, https://nyti.ms/2oiin5R.

[Meehl1954] P. E. Meehl, *Clinical Versus Statistical Prediction*, Brattleboro, VT: Echo Point Books & Media, LLC, 2013 (Originally: University of Minnesota, 1954), https://amzn.to/2BWWu5V.

[Kahneman2011] D. Kahneman, *Thinking Fast and Slow*, New York: MacMillan, 2011, https://amzn.to/2BZW9PT.

[Wainer2015] H. Wainer and R. Feinberg, "For want of a nail: Why unnecessarily long tests may be impeding the progress of Western civilization," *Significance*, The Royal Statistical Society, February 2015, http://ibm.biz/RSS-Significance-TestEfficiency.

[Levine2002] M. Levine, *One Mind at a Time*, New York: Simon and Schuster, 2002, https://amzn.to/2PJdQ9f.

[Mislevy2003a] R. J. Mislevy and R Zwick, "Scaling, linking, and reporting in a periodic assessment," *Journal of Education Measurement*, National Council on Measurements in Education, Summer 2002, http://ibm.biz/NCME-TCSA.

[DoEd2016] "Issue brief: Early warning systems," Policy and Programs Studies Service; Office of Planning, Evaluation, and Policy Development; U.S. Department of Education, September 2016, http://ibm.biz/DoEd-EWS.