

Reflections on Positioning Diagnostic Classification Modeling in Classroom Assessment

Justin Paulsen | jumpaul@indiana.edu
Christopher D. Andrews | andrewch@indiana.edu
Indiana University - Bloomington

Cognitive diagnostic model (CDMs) are of great interest in the measurement community because of their potential to model learning theories using multiple latent constructs that would allow more fine-grained assessment of students' strengths and weaknesses (Bradshaw, 2017). CDMs are ideally based on assessments designed to elicit specific processes or knowledge structures demonstrated by cognitive science or learning science to be necessary for responding correctly to the given item (Nichols, 1994). A simple example of this from the literature comes from fraction subtraction where solving an individual item requires multiple, discrete processes such as converting a whole number to a fraction, separating the whole number from a fraction, simplifying, and subtracting numerators (Tatsuoka, 1990). Constructing a cognitive diagnostic assessment (CDA) like this one that effectively balances these fraction subtraction skills requires a well-developed theory about the cognitive processes involved, which is difficult to develop and uncommon in the literature (Bradshaw, 2017). However, such a test fitted with a CDM could identify which of the skills each student does and does not possess. These classifications are typically binary (e.g. mastery v. non-mastery), but may also be categorical (e.g. proficiency levels). In this example, a student's responses might indicate that he or she has mastered converting a whole number to a fraction but has not mastered simplifying the fraction. Such information is obviously different from traditional assessments that provide total scores or percentiles that primarily serve to rank students as it can guide learning and teaching decision making at the individual level (Nicholas, 1994; Rupp, Templin, Henson, 2010).

The fine-grained, attribute classifications in CDMs are well suited to formative assessment. Formative assessment is defined here as the use of assessment outcomes to guide teaching and learning decisions (Black & Wiliam, 1998; Black, Wilson, & Yao, 2011). As noted above, CDMs consider a general construct at the level of cognitive processes and knowledge structures which helps identify what can be done to improve and guide further learning (Rupp, Templin, & Henson, 2010). Additionally, CDMs are more likely to explain why students perform as they do since the CDA items are explicitly built to elicit particular processes and knowledge structures whereas traditional measurement items are selected based on their relevance to the content domain (Leighton & Gierl, 2007; Nichols, 1994). Moreover, CDMs focus on mastery which motivates the student to achieve competency rather than a performative focus on grades or stanardized percentiles (Shepard, Penuel, & Pellegrino, 2018).

Formative assessment is most effective where the proximity of the assessment feedback is close to the instructional context (Hickey & Pellegrino, 2005; Shepard, Penuel, & Pellegrino, 2018). Proximity of assessment refers to the degree to which the assessment aligns with classroom practices and purposes (Hickey & Pellegrino, 2005). For example, a large-scale assessment is designed to be unrelated to a particular curriculum and thus would struggle to closely relate to a particular classroom environment (Hickey & Pellegrino, 2005). Empirical applications of CDMs have primarily focused on large-scale assessments (Sessoms & Henson, 2018). While CDMs have the potential to guide learning and instruction at an individual level, the literature has not considered how they would need to be adapted in order to align with

classroom practices to achieve this goal. In this paper, we raise questions about how CDMs would fit into the curriculum, instruction, and assessment system of the classroom; identify potential challenges to this scenario; and suggest possible solutions.

Before identifying these challenges, we describe the classroom context in which we develop our hypothetical implementation. We follow the suggestions of Shepard, Penuel, and Pellegrino (2018) and Wiliam (2018) that the sociocultural theory of learning best accounts for the many factors that influencing learning. At the classroom level, this means that learning takes the shape of “purposeful activity [that leads] to increasingly proficient contributions to a community of practice without having to be separately or artificially induced.” (Shepard, Penuel, & Pellegrino, p.23) This means that purposeful activity, without considerations of motivation and identity, is insufficient for a learner to develop into a meaningful contributor. Thus, assessment practices should orient toward displaying holistic student thinking rather than piecemeal evaluation, intrinsic motivation toward mastery rather than extrinsic motivation toward grades or teacher pleasing, and internalization of judgment criteria so that the student can deepen their understanding of the discipline (Shepard, Hammerness, Darling-Hammon, & Rust, 2005; Shepard, Penuel, Pellegrino, 2018). Assessment must align with classroom instructional practices so as to ensure they reinforce the student’s development.

Based on this review, we identified three key questions highlighting the disconnect between the potential for CDMs to provide diagnostic feedback to be used to guide teaching and learning and the needs of a classroom in its assessment.

1. How can CDM classifications fit into the existing classroom assessment system in a way that does not distort instructional or assessment practices?
2. How do we maintain focus on quality of learning when CDM classifications report discrete classifications?
3. How can CDM classifications provide more guidance to students on next steps and encourage student self-regulation as Shepard, Penuel, and Pellegrino (2018) suggest?

Question 1: CDM Classifications and Distortion of Classroom Practices

Implementing any assessment into the classroom has the potential to distort its instructional and assessment practices because the assessment signals to teachers and students what is valued. This can result in narrowing curriculum, reducing the focus on helping students develop their identities as practitioners in a community. CDMs may aggravate this distortion in the following ways. First, CDAs typically will have to be developed by non-teachers (e.g. test developers, psychometricians) creating distance between in-class experiences and assessment, removing control from the teacher, and jeopardizing the consistency of criteria for evaluating work across assessment instances. Second, CDMs are a large sample model which make it difficult to conform to the particular curriculum and instruction being used in a class or a school. Third, outcomes must be interpretable and usable by a teacher so that it links into the instructional decisions, in a way that doesn't add more work for the teacher. Relatedly, CDM classification introduces complex data that has to be effectively analyzed, interpreted, and integrated into a teaching plan.

Question 2: CDM Classifications and Focus on Quality of Learning

A concern among adherents to the sociocultural theory of learning is that quantifying learning in formative assessment works counter to the desire to help students develop as contributors and participants. Formative assessment that is more aligned with this process of learning would be individual or group conversations that demonstrate student learning and allow teachers to provide feedback contextualized to the moment. CDM classifications present learning

as unequivocal. Learning is discretized into components that are mastered or not mastered, without really capturing the student's holistic sense of development within the learning community. Similarly, CDM classifications typically do not provide descriptive features, thus not providing teachers with observation of the task. This effectively alienates the teacher from understanding the underlying cognitive process and being able to contextualize the CDM classification. The teacher could effectively be removed from the learning process as the classifications themselves are signals to the student, and the teachers has relatively little control over the assessment process and results.

Question 3: CDM Classifications and Student Self-Regulation

Research around formative feedback has demonstrated that it can lead to significant gains in learning, and this is most likely to occur when feedback illustrates to students how they can individually improve (Black & Wiliam, 1998). Such feedback must indicate more about the quality of the learning, contextualizing results in terms of the details of the learning process; scores alone are insufficient to signal how to develop mastery (Shepard, Penuel, Pellegrino, 2018). CDM classifications do provide more detailed score reports by reporting on multiple cognitive processes or knowledge structures. However, it still provides little direction on what students should do based on their results. If our goal is to help students develop further in the given community of practice, CDA outcomes may point toward a more discrete, rather than holistic vision. Additionally, providing feedback is a highly emotional interactive process, and CDA's outcome orientation towards mastery or proficiency levels may result in more affective than cognitive impacts if not approached appropriately.

Potential Solutions

Many of the issues addressed above cannot be explicitly addressed, but we believe that certain steps can be taken to encourage CDM implementation in the classroom that better meets the needs of students and teachers.

- Measurement experts should identify diagnostic classification methods that function reliably and validly in the context of small samples. Such small sample methods could allow for modeling CDAs adapted specifically for the curriculum and instructional approaches of particular teachers and schools. This would also enable teachers to be participants in the development process, empowering them to design and use the classifications to meet their needs.
- CDAs need to be designed to provide granular feedback that is actionable. A review of empirical applications of CDM classifications indicates that most consider attributes at a higher, abstract level (e.g. algebra, geometry, proportional reasoning; or vocabulary, syntax, connecting and synthesizing) (see Sessoms & Henson, 2018). Classifications on such general attributes cannot speak to quality of learning or evoke student self-regulation. More detailed, specific attributes should be measured (e.g. solving a word problem using equivalent ratios, identifying ratios from word problems (Gierl, Majeau, Alves, 2010)) can provide the kind of feedback students and teachers can incorporate in the classroom.
- CDM classifications should be accompanied by detailed contextualization of the meaning of the classifications. A good example of this is American Council on the Teaching of Foreign Languages "Can-Do" statements that accompany classifications given through the Oral Proficiency Interview. These statements provide concrete evidences about how a person with a given classification will speak along a variety of grammatical issues. They also provide descriptors for the classification just above so that learners clearly

understand what skills would be necessary for advancing to the next level, encouraging self-regulation. Such contextualization of classifications help teachers recognize how the assessment results link to classroom learning behaviors and enable them to guide students in their learning, while also motivating students by providing them a vision of how to develop mastery.

By recommending these solutions, we do not suggest that these are simple and unchallenging but rather that they are necessary in order to better fit into the classroom ecology.

References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective*, 9(2-3), 71-123.
- Bradshaw, L. (2017). Diagnostic classification models. In Rupp, A. & Leighton, J. (eds.) *The handbook of cognition and assessment, first edition*. Oxford: Wiley Blackwell.
- Gierl, M., Alves, C., & Majeau, R. (2010). Using the attribute hierarchy method to make diagnostic inferences about examinees' knowledge and skills in mathematics: An operational implementation of cognitive diagnostic assessment. *International Journal of Testing*, 10(4), 318-341.
- Hickey, D., & Pellegrino, J. W. (2005). Theory, level, and function. *Transfer of learning from a modern multidisciplinary perspective*, 251-293.
- Leighton, J. P., & Gierl, M. J. (2007). Why cognitive diagnostic assessment. *Cognitive diagnostic assessment for education: Theory and applications*, 3-18.
- Nichols, P. D. (1994). A framework for developing cognitively diagnostic assessments. *Review of Educational Research*, 64(4), 575-603.
- Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications*. Guilford Press.
- Sessoms, J., & Henson, R. A. (2018). Applications of Diagnostic Classification Models: A Literature Review and Critical Commentary. *Measurement: Interdisciplinary Research and Perspectives*, 16(1), 1-17.
- Shepard, L., Hammerness, K., Darling-Hammond, L., & Rust, F. (2005). Assessment. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 275–326). San Francisco, CA: Jossey-Bass.
- Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, 37(1), 21-34.
- Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. *Diagnostic monitoring of skill and knowledge acquisition*, 453-488.
- Wiliam, D. (2018). How can assessment support learning? A response to Wilson and Shepard, Penuel, and Pellegrino. *Educational Measurement: Issues and Practices*, 37(1), 42-44.