Research Report

A Preliminary Validity Evaluation of a Learning Progression for the Concept of Function

ETS RR–19-21

Edith Aurora Graf
Stephanie Peters
James H. Fife
Peter W. van Rijn
Meirav Arieli-Attali
Elizabeth Marquez

December 2019

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

ETS Research Report Series ISSN 2330-8516

RESEARCH REPORT

A Preliminary Validity Evaluation of a Learning Progression for the Concept of Function

Edith Aurora Graf,[1] Stephanie Peters,[1] James H. Fife,[1] Peter W. van Rijn,[2] Meirav Arieli-Attali,[3] & Elizabeth Marquez[1]

1 Educational Testing Service, Princeton, NJ
2 ETS Global, Amsterdam, The Netherlands
3 Fordham University, New York, NY

Learning progressions (LPs) describe the development of domain-specific knowledge, skills, and understanding. Each level of an LP characterizes a phase of student thinking en route to a target performance. The rationale behind LP development is to provide road maps that can be used to guide student thinking from one level to the next. The validity of an LP cannot be taken for granted, however. LPs evolve from a synthesis of multiple research studies, subject-matter expertise, and standards documents. They are working models of student development that may require revision in light of critique and empirical evidence. The formulation of an LP is an iterative process in which expert feedback is elicited, data are collected, and the LP is revised accordingly. We developed an LP for the concept of function both because the concept of function is challenging to attain and because it is central to the study of algebra and higher mathematics. We report early findings with respect to the validity of the concept of function LP, based on small-scale cognitive interviews and expert reviews.

Keywords Learning progressions; mathematics assessment; function concept

doi:10.1002/ets2.12257

This report is organized into several main sections. In the opening section, we discuss what learning progressions (LPs) are and why it is important to validate the interpretations and decisions based on them. We outline a number of steps in constructing a validity argument and the questions addressed by each of these steps. We provide the research questions that guided this work, namely, (a) How well does student thinking elicited from cognitive interviews align with a proposed LP for the concept of function? (b) What revisions to the LP and the tasks based on it are suggested by subject-matter experts? and (c) What can we learn from the way experts rate student responses using the levels of the LP? In the next section, we provide the rationale for developing an LP for the concept of function and summarize earlier work in this area. Following an overview of the steps taken to develop the progression, we present the LP itself.

In later sections, we discuss procedures for conducting cognitive interviews and eliciting feedback and scores from an expert panel. We present findings from three data sets: The first data set consists of cognitive interviews conducted with students working on tasks based on the LP, the second data set consists of feedback on the LP from a panel of experts, and the third consists of student responses scored with respect to the LP by the same panel of experts. The analyses of the first two data sets are primarily qualitative, while the analysis of the third data set is primarily quantitative. We conclude with a discussion of the findings and how we used them to produce a revised version of the LP.

Learning Progressions and Validation

LPs describe the development of domain-specific knowledge, skills, and understanding. The term *learning progression* originated in science education (e.g., Smith, Wiser, Anderson, Krajcik, & Coppola, 2004); its close sibling, the *learning trajectory*, has its roots in mathematics education (e.g., Clements & Sarama, 2004; Confrey, Maloney, Nguyen, Mojica, & Myers, 2009; Simon, 1995). We use the terms interchangeably here.

In our work, we draw on the LP definition from the *CBAL*® research initiative (Bennett, 2010; Bennett & Gitomer, 2009). In CBAL, an LP is characterized as

Corresponding author: E. A. Graf, E-mail: agraf@ets.org

> a description of qualitative change in a student's level of sophistication for a key concept, process, strategy, practice or habit of mind. Change may occur due to a variety of factors, including maturation and instruction, and each progression is presumed to hold for most, but not all, students. As with all scientific research, the progressions are open to empirical verification and theoretical challenge. (Educational Testing Service, n.d., list item #2)

An LP may be used to guide instructional decisions, pending empirical support. It is important to note here that "empirically supported" is a different requirement from "research based." Ideally, LPs are research based, that is, grounded in literature pertaining to theory, item difficulty factors, or case studies that suggest how and when big ideas emerge. A research foundation is a necessary, but not sufficient, requirement for an LP. Usually, LPs are based on research as well as on an analysis of the logical structure of the domain (Daro, Mosher, & Corcoran, 2011). Expert opinion is used to modify the LP and fill in the gaps. Each of the research studies on which an LP is based typically addresses one concept or problem type and samples from a particular population (e.g., ninth graders). But an LP usually addresses a number of concepts that are assumed to develop across years. The development of an LP involves piecing together learning goals from standards documents, results from multiple studies, and expert opinion. It therefore constitutes a new theory that must be independently verified.

Kane (2006) distinguished between two types of arguments in validation, an *interpretive argument* and a *validity argument*. According to Kane, "an interpretive argument specifies the proposed interpretations and uses of test results by laying out the network of inferences and assumptions leading from the observed performances to the conclusions and decisions based on the performances," and "the validity argument provides an evaluation of the interpretive argument (Cronbach, 1988)" (p. 23). The interpretive argument is referred to as the interpretation/use argument (IUA) in Kane's more recent work (e.g., Kane, 2013), to reflect the emphasis on test-based decisions as well as interpretations. Kane and Bejar (2014) wrote,

> The IUA for assessment based on a learning progression would start with the student performances on the assessment tasks and would end with conclusions about the student (e.g., where the student is in the learning progression), and in applied settings, with suggestions about what to do next. (p. 120)

Because an LP characterizes the performances typical of different levels of student thinking, it provides proposed interpretations. If, in addition, an LP provides guidance for instruction based on those interpretations, it provides the use component of an IUA as well. When we refer to a "validity argument for an LP," we make the assumption that the LP guides both interpretation and instructional decisions based on student performances.

Kane (2006, 2013) characterized kinds of inferences that are commonly included in an IUA, such as *scoring*, *generalization*, *extrapolation*, and *decision*. "A *scoring inference* takes us from the observed performances to the observed score" (Kane, 2013, p. 10). A *generalization inference* is made from "the observed sample of performances to claims about expected performance in a universe of possible observations (most of which were not made) or to an estimated trait value that can be used to draw conclusions about the future performances" (Kane, 2013, p. 10). *Extrapolation inferences* "extend the interpretation into new performance domains," (Kane, 2013, p. 11), and a *decision inference* "takes us from a person's score to a decision about the person (or about an educational program or teacher)" (Kane, 2013, p. 11).

How are the inferences described by Kane (2006, 2013) operationalized in the development of an IUA for an LP? This is the focus of Kane and Bejar's (2014) article and is discussed here. To assess student standing with respect to an LP, typically an assessment is developed that is intended to elicit evidence associated with the different levels. As noted by van Rijn, Graf, Arieli-Attali, and Song (2018, p. 2), "LP levels can be assigned to performance at different grain sizes of the assessment." For example, a level can be assigned holistically to a complex performance or collectively to performances across several items. Another approach is to assign a level to each item so that each correct item response indicates reasoning at or above a particular level and each incorrect item response indicates reasoning somewhere below a particular level. Yet another approach is to assign levels to responses, where different responses to the same item may be associated with different levels of understanding. All of these methods involve assigning an observed score to an observed performance, and hence all are examples of *scoring inferences*. Students can be classified into LP levels based on cut scores derived from item response theory (IRT) models (e.g., Graf & van Rijn, 2016; van Rijn, Graf, & Deane, 2014); doing so involves a *generalization inference*. A prediction about the level of thinking a student will show on a classroom-based task involves an *extrapolation inference*, and an instructional next step based on a student's LP classification involves a *decision inference*.

Constructing a validity argument for an LP may be supported by a cycle of several steps (Graf & van Rijn, 2016). The first step involves theory development, or domain analysis (Mislevy, Steinberg, & Almond, 2003; Riconscente, Mislevy, & Corrigan, 2016). The next step involves empirical recovery of the levels of the LP. For example, an LP expressed as levels assumes that the levels are distinct and, in general, that they develop in a particular sequence (though some slippage between levels is expected). Both of these assumptions can be examined with cross-sectional and/or longitudinal data, using a variety of psychometric methods (such as IRT). If the analysis shows that the levels are distinct and ordered as expressed by the LP, then the levels are said to be "empirically recovered." The third step involves a comparison to competing models (because other learning theories might also account for the specification and ordering of levels). The last step involves an evaluation of instructional efficacy, that is, does instruction that is based on the LP yield better outcomes than traditional instruction? The cycle may include subcycles—for example, there may be a return to theory development following an evaluation of empirical recovery. We do not intend to suggest that the LP must go through the entire cycle before it can be used as a guide for assessment or instructional decisions—to the contrary, completing the validation cycle requires that it be used for both. We do suggest, however, that the LP be used provisionally, pending at least some empirical support. In some sense, an LP is always provisional, because it incorporates expectations about the nature of instruction (National Assessment Governing Board, 2008), and the nature of instruction is subject to change.

The work described here is a preliminary validity evaluation of an LP for the concept of function (CoF). As such, it takes place between the analysis of the domain and the empirical recovery at scale. Arieli-Attali and Cayton-Hodges (2014) suggested using the results from cognitive interviews with students to refine an LP for rational numbers. We used this procedure as an initial step in the process of constructing a validity argument for the CoF and the tasks designed from it. It is important to note that for the collection of cognitive interviews, we used samples of convenience. Following this, we asked experts to provide written feedback on the CoF LP, the tasks, and sample student responses to the tasks. We also asked experts to classify student responses into levels of the LP. Agreement among teachers with respect to classification of student responses into levels of an LP has been examined by van Rijn et al. (2018). It has been argued that placing a student into a particular level of an LP is not necessarily an appropriate task, because his or her performance may show aspects of multiple levels (Battista, 2011; Corcoran, Mosher, & Rogat, 2009; Daro et al., 2011). For example, in a discussion of the hierarchic interactionalism framework of Clements and Sarama, Daro et al. (2011) wrote the following with respect to placing students into levels of an LP:

> The levels are seen as being qualitatively distinct cognitive structures of "increasing sophistication, complexity, abstraction, power, and generality." For the most part they are thought to develop gradually out of the preceding level(s) rather than being sudden reconfigurations, and that means that students often can be considered to be partially at one level while showing some of the characteristics of the next, and "placing" them in order to assign challenging, but doable work becomes a matter of making probabilistic judgments that they are more likely to perform in ways characteristic of a particular level than those of levels that come before or after it. (p. 24)

Therefore, for each response, we asked experts to provide the probability that the response was at each level.

We have three main research questions, the third of which includes two subquestions:

1. Based on the results from the cognitive interviews, what student thinking is elicited by tasks designed to assess understanding of the CoF, and how well does this align with the LP?
2. What revisions to the CoF LP and the tasks are suggested by subject-matter experts?
3. What can we learn from the way experts rate student responses using the levels of the LP?
   a. What are the features of student responses that are not classified consistently among raters?
   b. What is the level of agreement among raters who are classifying student responses into levels of the LP?

The answers to any of these questions might suggest revisions to the tasks, the LPs, or both.

The Concept of Function Learning Progression

The CoF is both central to the study of mathematics and challenging for students to learn. In our earlier work for CBAL mathematics, in which we reviewed the literature in cognitive psychology and mathematics education (Graf, 2009; Graf, Harris, Marquez, Fife, & Redman, 2009, 2010), we identified working with functions as an important skill. The Common

Core State Standards for Mathematics (CCSSM; Common Core State Standards Initiative, 2010) emphasize the importance of the CoF. For example, eighth-grade Standard 8.F includes the text “Understand that a function is a rule that assigns to each input exactly one output. The graph of a function is the set of ordered pairs consisting of an input and the corresponding output,” and high school Standard F-IF includes “Understand that a function from one set (called the domain) to another set (called the range) assigns to each element of the domain exactly one element of the range. If f is a function and x is an element of its domain, then $f(x)$ denotes the output of f corresponding to the input x. The graph of f is the graph of the equation $y = f(x)$.” The ordering of these standards is consistent with the research, namely, understanding of an operational notion of function (as expressed by Standard 8.F) precedes understanding of the formal, set-theoretic definition (as expressed by Standard F-IF). Vinner and Dreyfus (1989) found that few college students have an understanding of the formal definition of function, and results from Sfard (1992) were consistent with this finding.

The CoF LP has several predecessors. These include Sfard’s (1991, 1992) model of concept development as applied to functions; APOS theory as applied to functions (e.g., Dubinsky & Harel, 1992; Dubinsky & Wilson, 2013); a model for understanding functions developed by Kalchman, Moss, and Case (2001) and elaborated in Kalchman and Koedinger (2005); an LP for functions developed by Wilmot, Schoenfeld, Wilson, Champney, and Zahner (2011); and an LP for functions by Arieli-Attali, Wylie, and Bauer (2012) and discussed in Graf and Arieli-Attali (2015).

Sfard’s (1991, 1992) model consists of three stages: (a) *interiorization*, (b) *condensation*, and (c) *reification*. During interiorization, a student becomes familiar with operational processes. Once the operational steps are well understood, attention shifts from discrete steps to the result. Finally, during reification, the concept is realized as a structural object that can support the development of new concepts. In the context of functions, Sfard (1992) argued that an operational notion of function is intuitive and should precede introduction of the formal, set-theoretic definition, which represents complete reification of the concept.

APOS theory (Dubinsky & Harel, 1992; Dubinsky & Wilson, 2013) is described with respect to the stages *action*, *process*, and *object*, which are enacted according to *schemas*. At the *prefunction* stage, no CoF has developed. At the *action* stage, the student can find outputs from inputs, one at a time. At the *process* stage, the student has internalized this procedure and can carry it out mentally. Finally, at the *object* stage, the student can operate on functions. As Dubinsky and colleagues described it, composing functions and finding inverses first happens at the process stage, and finding derivatives happens at the object stage. A student may demonstrate an approach that falls at or below the highest stage he or she has attained, depending on the demands of the task.

In the four-level model of Kalchman et al. (2001), students can extend patterns at Level 1. By Level 2, they can generate ordered pairs and can plot bar and line graphs. At Level 3, they recognize families of linear and nonlinear functions, and at Level 4, they work in all four quadrants of the Cartesian coordinate system. In the six-level model of Wilmot et al. (2011), students become increasingly facile with translating among equivalent representations of functions. At the highest level, they are able to solve nonroutine problems in real-world contexts. The five levels of the LP of Arieli-Attali et al. (2012) include one-dimensional change (Level 1), mutual change (Level 2), constant change (Level 3), comparing rates of change (Level 4), and changing change (Level 5). Level 1 is equivalent to the first level of Kalchman et al. (2001). Level 2 involves being able to detect the direction, if not the magnitude, of change. At Level 3, students can work with linear functions, and at Level 4, they can compare slopes. Finally, at Level 5, students work with nonlinear functions, such as polynomials.

Although all of these models of development are for functions, they focus on different aspects of understanding. Sfard’s (1991, 1992) model and the APOS theory model (Dubinsky & Harel, 1992; Dubinsky & Wilson, 2013) emphasize understanding the properties of a relation that determine whether or not it is a function. The model of Arieli-Attali et al. (2012) emphasizes the concept of change. The models of Kalchman et al. (2001) and Wilmot et al. (2011) heavily emphasize representational fluency; the Kalchman et al. model also addresses the idea of function families. A common theme to all of these models is that they address representational fluency. The CoF LP draws on all of these models as well as other research findings but shares with the models of Sfard (1991, 1992) and Dubinsky and colleagues (Dubinsky & Harel, 1992; Dubinsky & Wilson, 2013) that the definition of function is an important component. In other words, the CoF LP integrates themes about the definition of function, representational fluency, and the concept of change.

As Plake, Huff, and Reshetar (2010) have noted, characterizing the achievement-level descriptors of an LP is an iterative process, and so it was with the development of the CoF LP. We went through several cycles of literature review, drafting achievement-level descriptors, and discussion. Fairly early on in our literature review, we recognized the presence of several themes, namely, the properties of a relation that determine whether it is a function, representational fluency, the concept of change, and function families. These themes helped to structure the levels—for example, representational fluency seemed to be a theme that crossed levels of understanding, while it seemed the concept of change could not logically coincide with a *pointwise* conception of function (Leinhardt, Zaslavsky, & Stein, 1990). The number of levels was determined by the number of qualitative transitions in understanding with respect to one or more themes as culled from the literature. For example, understanding the notion of dependence (that the values of one variable depend on the values of another) is a transition identified in the literature that served as one of the differentiators between Level 1 and Level 2. As another example, understanding the uniqueness property of a function indicates understanding at Level 4 (Ponce, 2007; Vinner & Dreyfus, 1989).

Though there were several relatively complete models of development for the CoF in the literature, there were also studies that focused on particular misconceptions or particular transitions in thinking. Development of the provisional LP was in part an exercise in weaving together partial orderings of understanding from the literature (e.g., if Study 1 suggests that Conception A precedes Conception B, and Study 2 suggests that Conception B precedes Conception C, one might propose that in the LP, Conception A precedes Conception B, which precedes Conception C). This weaving together of findings entails major assumptions, however, since the samples from different studies typically generalize to different populations, and the studies are typically carried out using different procedures.

We had quite a bit of discussion about Levels 5 and 6 and whether they should really be distinct. Eventually, we decided that a complete understanding of domain and range is an important qualitative transition to Level 6, and our read of the literature suggests that this is challenging to develop (Markovits, Eylon, & Bruckheimer, 1986).

In what follows, we discuss the nature of student thinking at different levels of the CoF LP.

Levels 1 and 2 (Preinstruction and Familiarization)

Consistent with the model of Kalchman et al. (2001), at Level 1 of the CoF LP, students can detect patterns in one variable by extending sequences but have not yet developed the notion of dependence (that the values of one variable depend on the values of the other). At Level 2, they have a limited notion of dependence (for example, they can recognize if a function is strictly increasing or decreasing), but their interpretation of graphs is primarily *pointwise*. Pointwise interpretation involves observing what happens at a point on the graph rather than attending to a trend (Leinhardt et al., 1990). Also at Level 2, students think of a function as a formula or computational process (Carlson & Oehrtman, 2005; Sfard, 1992). Students at this level will identify different formulas that represent the same function as different functions. For example, $y = 3x$ and $u = 3v$ might be considered different functions. Because they associate functions with formulas, students at this level are not yet translating among equivalent representations of a function—for example, they may not recognize that a graph and a formula represent the same function and would probably assert that a function represented by a graph is not a function.

A couple of difficulties emerge at Level 2 that persist through Level 4. One is confusion with pictorial aspects of the situation (Carlson & Oehrtman, 2005; Dugdale, 1993; Monk, 1992). One way this can happen is if a student interprets position as speed when looking at a graph. For example, consider the graph of Wanda's Walk in Figure 2. A student who interprets position as speed might say that Wanda is moving fastest between Points D and E. Another difficulty concerns overgeneralization of linearity (Carlson & Oehrtman, 2005; Karplus, 1979; Leinhardt et al., 1990). An example of this is if a student draws a straight line as the only function that will pass through a pair of points.

Level 3 (Making Connections)

At this level, the notion of dependence is developing (students think in terms of inputs and outputs), and students are close to the Grade 8 standard from the CCSSM, except that the *one-valuedness* idea (Ponce, 2007; Vinner & Dreyfus, 1989) that each input maps to exactly one output is not yet part of students' schemas. Students at this level are starting to translate among alternative representations of functions, though they are more likely to successfully translate from equations to tables and graphs than from tables or graphs to equations, since the latter is more difficult (Leinhardt et al., 1990; Markovits et al., 1986). Certain kinds of functions are less likely to be identified as such, including many-to-one functions (in particular constant functions), piecewise functions, discontinuous functions, and other functions that do not have an obvious rule or pattern (Leinhardt et al., 1990; Markovits et al., 1986). When presented with a piecewise function, students at this level may indicate that it is not a single function but several.

Level 4 (Synthesis)

By Level 4, students have developed the one-valuedness idea (Vinner & Dreyfus, 1989). They are beginning to translate from graphs and tables to equations. Because they have a fully developed notion of dependence and inputs and outputs, they are able to compose functions and find inverses (Carlson & Oehrtman, 2005). Students at this level have acquired a *process* view of functions (Dubinsky & Harel, 1992), in which they consider what is happening to the function as a whole rather than at single points. According to Carlson and Oehrtman (2005),

> a student with a process view can conceive of the entire process as happening to all values at once, and is able to conceptually run through a continuum of input values while attending to the resulting impact on output values. (para. 7)

This, Carlson and Oehrtman argued, is the skill needed to apply covariational reasoning (Confrey & Smith, 1994, 1995), where the student can visualize how one variable changes (or does not change) with the other. Covariational reasoning develops at Level 4.

Levels 5 and 6 (Thorough Conceptualization and Drawing Extensions)

At Level 5, students have developed a thorough conceptualization of the CoF as addressed in the Common Core high school Standard F-IF. They can translate among equivalent representations. They can transform functions and are starting to recognize the role of parameters in some function families, which is among the more difficult tasks for students to do (Kieran, 1993). They may still have difficulty in certain situations (e.g., recognizing discontinuous or nondeterministic functions as such). The concepts of domain and range are not necessarily completely understood. For example, when evaluating whether two functions are equivalent, the student may not consider the domain and range, focusing only on the rule that relates the two variables (Markovits et al., 1986). Level 6 is included for completeness—it represents a professional mathematician's CoF. An overview of the CoF LP is shown in Table 1.

Themes Addressed by the Concept of Function Learning Progression

As evident from the level descriptions, several main themes, or progress variables, as they are sometimes called, are addressed by the CoF LP. These include the following: (a) use of multiple representations, (b) understanding of the definition of function, (c) understanding covariation, and (d) understanding of functions as families. A summary of these themes and how they evolve across levels of the LP is given in Figure 1. The level descriptions also refer to emerging skills and difficulties. A timeline of concepts and skills is given in Table 2, and a timeline of difficulties is given in Table 3. An "X" in the timeline of concepts and skills means that students have that skill or understand that concept at a particular level. An "X" in the timeline of difficulties means that the preconception or misconception is still present at a particular level. The full CoF LP is given in Appendix A.

Cognitive Interviews

Method

Participants

Fifteen students in Grades 9–12 participated in the cognitive interviews. They were recruited through an online posting. There were six male and nine female participants. Four of the participants were in ninth grade, three were in 10th grade, five were in 11th grade, and three were in 12th grade. They were primarily from New Jersey and Pennsylvania school districts and enrolled in courses such as Algebra, Algebra I, Algebra I (Honors), Algebra II, Geometry, Precalculus, and Precalculus (With Limits). Students in this sample reported a minimum grade of B+ in mathematics.

Table 1 Concept of Function Learning Progression Overview

Level	Description	Characteristic	Schema
6	Drawing Extensions	Function families	Families of functions are perceived as parameterized objects, and the role of the parameters is understood; the role of domain and range is fully recognized.
5	Thorough Conceptualization	Function as an object	A function is perceived as an object that can be operated upon.
4	Synthesis	Covariational reasoning	The *one-valuedness* idea has developed, and students are facile in their use of alternate representations of functions. Students consider how variables covary and attend to global features of graphs.
3	Making Connections	Function as a rule	The concept of *dependence* is beginning to develop, but the notion of *one-valuedness* (that each input is mapped to exactly one output) is not firmly in place. There is some recognition that that a function can be captured using different representations. A function that does not appear to have a consistent "rule" throughout the domain may be rejected as a function.
2	Familiarization	Function as formula	A function is perceived as an algebraic formula or equation. In the student's view, different algebraic equations cannot represent the same function, even if they are equivalent.
1	Preinstruction	No function concept	A concept of function has not yet developed, but students can extend sequences.

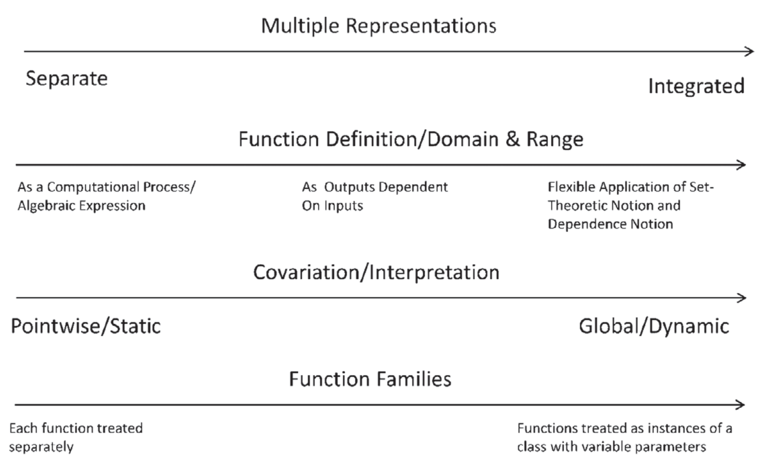


Figure 1 Themes addressed by the concept of function learning progression.

Materials

Because we did not want students' performance to be dependent on having memorized the definition of function, we provided them with definitions prior to solving the tasks. These definitions were adapted from those found in Collins et al. (2001, p. 802) and are as follows:

> function: 1. A relationship between input and output in which every input is paired with exactly one output. 2. A relation in which each element of the domain is paired with exactly one element of the range.

Three tasks designed to assess students' standing with respect to the CoF LP were developed: Wanda's Walk (WW; Marquez, 2015), Annika's Bakery (AB; Graf, 2015a), and Secret Messages (SM; Graf, 2015b). The tasks were developed to elicit evidence addressed by the levels of the LP. For example, WW was designed to focus on whether students can disentangle the meaning of a graph from its pictorial features and whether they can identify a many-to-one function as such. Task

Table 2 Timeline of Concepts and Skills

Skill/concept	Level 1	Level 2	Level 3	Level 4
Extend patterns in one dimension	X	X	X	X
Evaluate formulas		X	X	X
Plot points on a graph		X	X	X
Detect increasing/decreasing trends		X	X	X
Interpret function notation			X	X
Understand notion of dependency			X	X
Understand domain as a set of inputs			X	X
Understand range as a set of outputs			X	X
Recognize equivalent representations			X	X
Understand covariational reasoning				X
Understand "one-valuedness"				X
Compose and find inverses				X

Table 3 Timeline of Difficulties

Misconception/preconception	Level 2	Level 3	Level 4
Function as formula	X		
Representations are separate	X		
Not a function			
Functions defined by a graph	X		
Many-to-one functions	X	X	
Piecewise functions	X	X	X
Discontinuous functions	X	X	X
Nondeterministic functions	X	X	X
Difficulty translating to equation	X	X	X
Confusion with pictorial aspects	X	X	X
Overgeneralization of linearity	X	X	X

AB was designed to assess representational fluency and whether the representation used influences a student's decision about whether a relation is a function. Like WW, it also features a many-to-one function. Task SM uses relations that are functions as well as relations that are not because they fail the one-valuedness criterion, with the goal of ascertaining if students understand the one-valuedness property of a function as well as its contextualized implications. The tasks went through several iterations of review and revision by the project team, who made comments and suggestions by e-mail. The three tasks are scenario based and consist of 25 items combined. They are designed for paper-and-pencil delivery and include items in constructed response, multiple choice, and graphing formats.

In WW, a graph of Wanda's distance from home with respect to time is displayed with both multiple choice and constructed response questions intended to assess students' interpretation of a graphical representation of a function (see Figure 2 for the graph from WW). The first questions ask students to consider during which intervals Wanda spends the most time, the least time, and where she has the greatest speed and the least speed. Finally, the student is asked whether the graph represents a function.

The first question in AB involves a verbal description of a situation, as follows:

Annika's bakery is closed on Monday. On Tuesday through Thursday, the bakery prepares 3 chocolate fudge cakes each day. On Friday through Sunday, the bakery prepares 5 chocolate fudge cakes each day.

a. Is the number of chocolate fudge cakes prepared each day a function of the day of the week? (Explain).

In subsequent questions, the student is asked to translate between tables, graphs, and arrow diagrams. After working with each representation, the student is asked to reconsider whether or not the number of chocolate fudge cakes prepared each day is a function of the day of the week.

SM is a more advanced task; a cipher is used that maps letters to new letters or symbols. The student is first provided with a table that maps plaintext to ciphertext (see Figure 3).

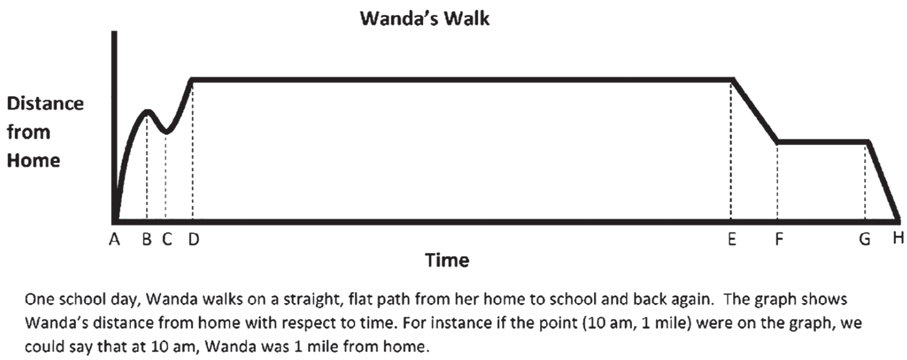


Figure 2 Graph of Wanda's Walk.

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	B	E	H	K	N	Q	T	W	Z	C	F	I	L	O	R	U	X	A	D	G	J	M	P	S	V	Y

Figure 3 The first table with which students are presented in Secret Messages.

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	E	G	I	K	M	O	Q	S	U	W	Y	A	C	E	G	I	K	M	O	Q	S	U	W	Y	A	C

Figure 4 The second table with which students are presented in Secret Messages.

The first several questions ask the student to encrypt and decrypt messages using the table. The next couple of questions ask the student to consider whether the mapping from plaintext to ciphertext is a function and whether the mapping from ciphertext to plaintext is a function. Next, the student is provided with a different table for mapping plaintext to ciphertext (see Figure 4).

The student is asked whether the rule in the table will work well for encrypting and decrypting messages, and then to determine whether the mapping from plaintext to ciphertext is a function, as well as vice versa. The complete tasks WW, AB, and SM are available in Appendix B.

Procedure

Cognitive interviews were conducted with one student at a time. Two researchers were present during each interview at all times: one in the role of interviewer and the other in the role of observer. Each session lasted 60 minutes, during which the student completed two to three CoF tasks designed to elicit evidence with respect to student standing on the CoF LP. The definitions of function were presented at the beginning of the session, and participants were free to refer to them throughout. Next, the tasks were presented. WW and AB were presented first, and time permitting, SM was presented last. The student first worked the solution on paper, and then explained the work. The interviewer asked the student to clarify explanations, providing prompts when necessary. To the extent possible, we used open-ended prompts to limit constraints on student responding. The interviewer asked a few general questions at the end, such as, "Which tasks did you prefer, and why?"

Results

In general, most students performed very well on these tasks. This was true even for SM, which was expected to be particularly difficult. Given that the minimum mathematics grade reported by students in this sample was B+ and that they

Table 4 Sample Responses From Wanda's Walk at Different Levels of the Learning Progression

Level	Explanation (learning progression interpretation)	Sample Response (does the graph of WW represent a function?)
4	At this level, students can skillfully use alternative equivalent representations.. .. [They] apply the **"one-valuedness" core idea in identifying functions (i.e., that each input is assigned to one and only one output).**	"Yes, it does represent a function. Looking closely at the graph shows that **each element of the domain (*x*-axis) is paired with exactly one element of the range (*y*-axis).** The vertical line test is a common way to test functions, and in this case the test worked. This graph is a representation of a function."
3	They [students] have an operational notion of function and are able to apply consistent **rules** to find outputs given inputs. In other words, a notion of dependence and outputs generated from inputs starts to develop at Level 3. **At this level, students would reject the following kinds of functions: (a) many-to-one functions,** (b) piecewise functions, (c) discontinuous functions, (d) functions with exceptional points, and (e) nondeterministic functions.	1. "No, depending on the distance she is from home being the input, it changes the time it takes for her to get back. Time = input, distance = output. **No, because there is no rule to the line.**" 2. "**Wanda's walk to and from school does not represent a function because she is at the same distance from her home at different times.** For instance, at both points A and H, she is at her house."
2	At this level, students can use an equation to generate ordered pairs in a table and/or plot points on a graph, but this is carried out as a sequence of procedures. However, students can grasp the **"trend"** in simple cases, namely, they can recognize when a dependent variable changes with an independent variable (the mutual change) and may be able to characterize it as strictly increasing or strictly decreasing.	"No, the distance away from home **gradually increases with time, but it also decreases.**"

Note. Boldfaced text in the Sample Response column are portions of the student response that correspond to boldfaced text in the Explanation (learning progression interpretation) column.. WW = Wanda's Walk.

volunteered to complete the CoF task set during the summer, it is likely that the sample consisted of highly motivated and/or high-achieving students. As noted, this was a sample of convenience recruited from an online posting—our findings might have been quite different had the sampling procedures been different.

Wanda's Walk

To answer WW completely correctly requires at least Level 4 thinking. However, initial evidence from the cognitive interviews suggests that the task may also elicit evidence of student thinking that corresponds to earlier levels. In answer to the question "At what point(s) in time is Wanda at home?" 13 participants selected A and H and two participants selected only A. All 15 participants selected D to E in response to the question "During which of the following time intervals does Wanda spend the most time?" All 15 participants also selected B to C in response to the question "During which of the following time intervals does Wanda spend the least time?" Fourteen participants selected A to B and one participant selected E to F in response to the question "During which of the following time intervals does Wanda have the greatest speed?"

In response to the question "Does the graph of Wanda's walk to and from school represent a function? Explain in terms of time and her distance from home," 11 participants responded "yes" and four participants responded "no." Sample student responses to this prompt are provided in Table 4, together with the levels of the LP with which they are aligned (as determined by the authors). For each student response, the parts that are relevant to the LP and the text from the LP to which they correspond are highlighted.

The responses in Table 4 are more or less canonical. The Level 4 response shows evidence of the *one-valuedness* concept (that each element of the domain is paired with exactly one element of the range). The first response categorized as Level 3 suggests that the student rejects the graph as a function because there is no rule underlying it. The second response categorized as a Level 3 suggests that the function is not a function because it is many-to-one (many time points map to the same position). In the Level 2 response, the student observes a trend (that the graph increases and then decreases) but rejects the function as such.

Annika's Bakery

For task AB, responses to the prompt to "make a graph that shows the number of chocolate fudge cakes prepared on each of the 7 days of the week" and to consider whether "the number of chocolate fudge cakes prepared each day [is] a function of the day of the week" are provided in the right column of Table 5. Note that the graphs were drawn on paper and pencil and have been rerendered in the figure. In contrast to the responses in Table 4, two of these responses (the second and the third) are less canonical, and hence more challenging to categorize using the levels of the LP. Instead of assigning a single level to each response, then, we assign a probability distribution to each response (based on the judgment of one of the authors), where what is assigned is the probability that the response is at each level (probabilities for each response sum to 1). These probabilities are shown in the far left column. The specific aspects of each level of the LP with which the responses are aligned are shown in the middle column. The distribution of probabilities reflects the uncertainty about the categorization of a response. For example, the first response in Table 5 is a benchmark Level 4, hence Level 4 is assigned with probability 1.00. Responses 2, 3, and 4 are much less canonical, and this is reflected in their distributions.

Results from AB suggest that one misconception may persist into higher levels than we originally thought. Connecting points with line segments or curves when the function is only defined at certain points, a difficulty hypothesized to emerge at Level 2, persisted for students who otherwise demonstrated Level 4 thinking. Nine of the 15 students created a graph either with points connected by line segments or line segments only.

Of the 15 participants, four at some point rejected the idea that the situation represents a function. Of the four participants who at some point rejected the situation as representing a function, three rejected it on the grounds that the relation is many-to-one. It is possible that these students have an overly restrictive definition of function as a one-to-one function or that they believe that one-to-one and one-to-many relations are functions but many-to-one are not. Three of the four students who rejected the graph in WW as showing a function also rejected the situation in AB as a function for at least one of the representations. Another interesting finding of note from AB was that three of the four students changed their minds about whether the situation describes a function, depending on the representation. Two of these three changed their minds from believing the situation does not represent a function to believing that it does, one after making the graph and the other after making the arrow diagram.

Secret Messages

Performance on SM was generally very high. All 15 students correctly responded to the prompts that asked them to encrypt and decrypt messages using the first chart in the task. In the first chart, the mapping rule from plaintext to ciphertext is one-to-one, so both the encryption rule and the decryption rule are functions. Thirteen of 15 students thought that both the encryption rule and the decryption rule in the first chart represent functions; two students answered the question for the encryption rule only (and responded that it is a function). In the second chart, the mapping rule from plaintext to ciphertext is many-to-one, so the encryption rule is a function but the decryption rule is not. Three students did not respond to prompts for the second chart due to time constraints.

Of the 12 students who provided complete responses, 10 responded "yes" to the question of whether the encryption and decryption rules for the first chart represent functions, "no" to the question of whether the second chart will work well for encrypting and decrypting messages, "yes" to the question of whether the encryption rule for the second chart represents a function, and "no" to the question of whether the decryption rule for the second chart represents a function (i.e., these 10 students responded correctly to these prompts). Of the students who responded incorrectly, one indicated that neither the encryption rule nor the decryption rule for the second chart is a function, and this student also responded "no" to whether the graph in WW represents a function and "no" to whether each of the representations in AB represents a function. With respect to the second chart, the other student wrote, "Although this chart is incorrect, it still does represent a function because there is an input and an output." On WW, this student correctly responded that the graph represents a function but identified the input as the speed at which Wanda is traveling. On AB, this student initially indicated that the number of chocolate fudge cakes is not a function of the day of the week (but indicated that it is a function after making the graph). Thus there is evidence that some difficulties persisted across tasks. However, two students who gave incorrect

Table 5 Sample Responses from Annika's Bakery at Different Levels of the Learning Progression

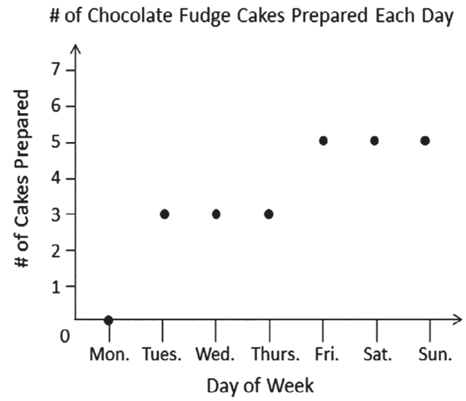
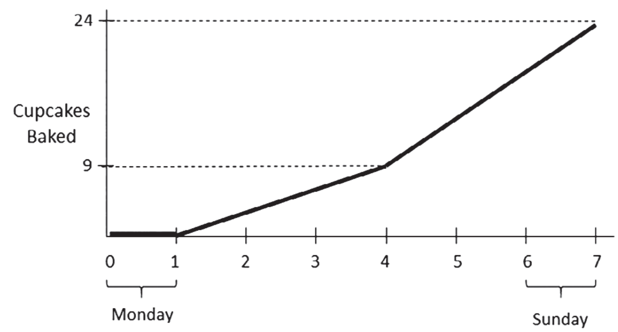
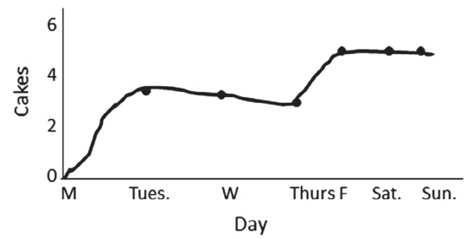
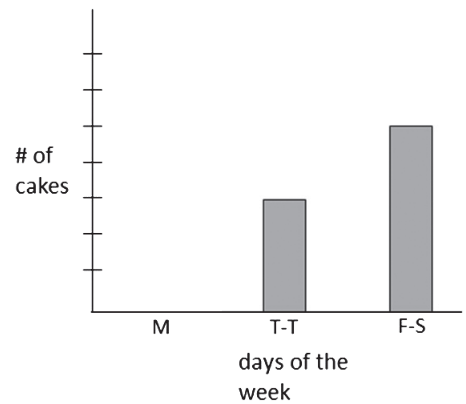
Level	Explanation (learning progression interpretation)	Sample response (make a graph that shows the number of chocolate fudge cakes prepared on each of the 7 days of the week)
6: 0.00 5: 0.00 4: 1.00 3: 0.00 2: 0.00 1: 0.00 0: 0.00	At this level, students can skillfully use alternative equivalent representations.. .. [They] apply the "one-valuedness" core idea in identifying functions (i.e., that each input is assigned to one and only one output). This student clearly has developed the "one-valuedness" concept. The points are not connected, and the axes are accurately labeled. The response is not above Level 4 because, according to the LP, this task would not *require* understanding above Level 4, and the student has not exceeded the requirements of the item.	 "The graph continues to support the fact that this is a function. Here, every element of the domain has been paired with only one element of the range. This means that the vertical line test passes and is still a function"
6: 0.00 5: 0.05 4: 0.20 3: 0.55 2: 0.20 1: 0.00 0: 0.00	This response was challenging to categorize because it answers a somewhat different question than what is asked (the graph is cumulative). Consistent with Level 2, the points are connected. But the response to some extent has exceeded the requirements of the item, because it suggests an interpretation of slope as the average number of cupcakes within an interval (Level 5). The student clearly has an operational notion of function and is able to apply a rule to find input given output (Level 3).	 "Number of cupcakes made during each day is represented by the slope of the graph. When no cupcakes are made, the graph is flat' where more cupcakes are made, the slope increases."
6: 0.00 5: 0.00 4: 0.35 3: 0.45 2: 0.20 1: 0.00 0: 0.00	The probabilities reflect the uncertainty in categorizing this response. It combines features of Levels 2, 3, and 4. Consistent with Level 2, the points are connected, but consistent with Level 3, the notion of dependence has developed. Consistent with Level 4, the notion of "one-valuedness" has developed.	 [Student made this graph when asked to make a table] "I still believe this [that it represents a function] because it shows me that the day of the week (the independent variable) appears to determine how many cakes are made (the dependent variable). I also believe it because the input is paired with exactly one output."

Table 5 Continued

Level	Explanation (learning progression interpretation)	Sample response (make a graph that shows the number of chocolate fudge cakes prepared on each of the 7 days of the week)
6: 0.00 5: 0.00 4: 0.00 3: 0.40 2: 0.60 1: 0.00 0: 0.00	This response reflects the misunderstanding that many-to-one functions are not functions, which is typical of Levels 2 and 3.	 "I still believe that the number of cakes is not a function of the day of the week. If you look at the graph, you can see the amount of cakes sometimes stays the same for different days of the week. If the number of cakes is the range and the day of the week is the domain, then some elements of the domain are paired with the same range."

responses on both WW and AB responded correctly to SM, so it is possible that SM was easier or that learning occurred during the course of the interview for these students.

Table 6 shows sample responses from three students to three prompts from SM. The prompts appear in the top row of the table, and responses from each student appear in each of the following rows. The first two sets of responses are clear Level 4s. The third set of responses is likely also Level 4, but since the responses are brief and less precise, this is more uncertain than for the first two sets of responses. If we were to assign probability distributions to these sets of responses, the first and second sets would receive a Level 4 with very high probability, but the probabilities in the third set would be more distributed across levels.

Expert Panel

Method

Panelists

The expert panel consisted of four subject-matter experts. The panelists were recruited based on recommendations from colleagues with whom they had done similar work. Two of the panelists are from universities where they hold professorships: one in mathematics and the other in mathematics education. The other two panelists are from the Assessment Development division of Educational Testing Services (ETS) in Princeton, New Jersey, and have expertise in item development and assessment design.

Materials

The panel was provided a packet of materials that included a brief summary of the literature on which the CoF is based, the full CoF LP, the CoF task set (WW, AB, and SM), sample student responses to the CoF task set (the samples were those in Table 4 and Table 5), and guiding questions to frame their reviews of the CoF LP. They were also provided a spreadsheet in which to complete a level-scoring activity, together with instructions for how to complete the activity. The guiding questions are as follows:

Table 6 Sample Responses From Secret Messages

Is the mapping from the plaintext to the ciphertext a function? Is the mapping from the ciphertext to the plaintext a function? Explain. (First chart)	Do you think this chart will work well for encrypting and decrypting messages? Why or why not? (Second chart)	Is the mapping from the plaintext to the ciphertext a function? Is the mapping from the ciphertext to the plaintext a function? Explain. (Second chart)
The mapping from plain text to ciphertext is a function, and the mapping from ciphertext to plain text is a function, because each letter of the alphabet (plaintext) corresponds to only one ciphertext letter. The same is to be said of the decryption of ciphertext to plaintext. Plaintext “A” corresponds to only ciphertext “B.”	No, the new chart has multiple letters that correspond to a ciphertext letter. Although encryption will be fine, the decryption is impossible. How do I know if ciphertext “A” corresponds to “L” or “Y”?	The mapping from plaintext to ciphertext is a function because each of the plaintext corresponds to only one ciphertext letter. The ciphertext to plaintext is not because each cipher text corresponds to more than one plaintext letter.
The mapping of the plaintext to the ciphertext and the mapping from the ciphertext to the plaintext are both functions, because any one element from the plaintext or the ciphertext is paired with one element from the other text.	This chart will not work for encrypting and decrypting messages because the ciphertext has repeated letters. So you wouldn’t be able to tell which letter from the plaintext it goes with.	The mapping from the plaintext to ciphertext is a function because each letter has one letter that it is paired with. The mapping from ciphertext to plaintext is not a function because certain letters are paired with two letters.
They are both functions, because both of them have only one.	It will not work well because there is more than one output for some inputs.	Plaintext to ciphertext is because no two letters are repeated in plaintext, but in ciphertext, it is not, because it does repeat letters.

1 Can this LP be used as a framework for designing assessment and research right now? Please note any internal consistencies in the LP.
2 Is the LP consistent with research or content from related work on LPs or learning trajectories?
3 Are the examples of student work consistent or inconsistent with the LP?
4 How do you think this is useful for the purpose of assessment development?
5 How do you think this is useful for the purpose of advancing the research field?
6 How do you think this is useful for the purpose of informing student learning and teacher instruction?
7 How can this be useful for the purpose of informing the construction of psychometric models?
8 To what extent is the LP aligned with the CCSSM?

The first few rows of the level-scoring spreadsheet are shown in Figure 5. The first column has student responses to the prompt “Does the graph of Wanda’s walk to and from school represent a function? Explain in terms of time and her distance from home” from WW. Responses from 11 students were included (the responses from the other 4 students, given in Table 4, were excluded since they were given as samples). The second through eighth columns correspond to score categories for each level of the LP. Level 0 is for responses that are missing, below Level 1, or unclear. Panelists were asked to enter the probability that a response is at each level, as we did for the responses in Table 5.

The ninth column of the spreadsheet shows the sum of the probabilities in the level columns. Initially, this column was highlighted in red with a sum of 0. As panelists entered probabilities for a response, the sum changed, and once the probabilities for the response summed to 1, the sum cell changed from red to white. Panelists entered their comments pertaining to how responses were scored in the final column. The first row was already complete when the spreadsheet was given to the panelists; it was intended as an example of how the spreadsheet should be filled out.

Procedure

Panelists completed their work over the course of about a month. They engaged in two main activities: providing written feedback on the LP and scoring student responses using levels of the LP. We used the same procedure with the panelists

Student Response	Level 0	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Sum	Comments
Response	0.25	0.60	0.12	0.03	0.00	0.00	0.00	1.00	It seemed to me that this response was actually much more sophisticated than what would be suggested by a Level 1 or Level 0, but the LP description suggested this was closest to Level 1, with some features of Level 0 and some features of Level 2.
Yes, her distance from home is a function of time, where distance is the dependent variable and time is the independent variable. At any one time she is only one distance from home.								0.00	
Yes, because for every input of time she has gone through a new output of distance from home is given.								0.00	

Figure 5 Level-scoring spreadsheet completed by each member of the expert panel.

from the universities and the panelists from ETS, however, the time frames for the two sets of panelists were staggered so that the procedure was conducted twice: once for the university panelists and once for the ETS panelists. Before they began their work, we gave each set of panelists a 45-minute slide presentation in which we introduced the materials and outlined the guiding questions, and discussed general expectations for the reviews and the level-scoring activity. Both introductory sessions were conducted remotely by video conference. The materials were e-mailed to panelists in advance of the introductory sessions so that they could ask questions about them at the sessions. They were asked to read the literature summary; to study the CoF LP, tasks, and student responses; and finally to do the level-scoring activity, time permitting. Each panelist was asked to deliver a synthesis of impressions about the LP, the research behind it, and the samples of student work and also to make specific suggestions for revisions. Although the guiding questions were provided, the task was deliberately open ended.

For the level-scoring activity, each panelist completed the spreadsheet independently. Panelists were reminded that there are no “correct” answers and that the probabilities should “reflect your judgment based on your interpretation of the LP.” They were encouraged to assign probabilities to levels based on how the LP is written, even when they disagreed with the level descriptors. For example, a response that seems very sophisticated but is still consistent with the LP level description for Level 2 should be assigned a high probability for Level 2. All panelists had sufficient time to complete the spreadsheet.

Each panelist returned his or her materials (the review and the completed spreadsheet) to us via e-mail. We reviewed the materials and developed slide presentations summarizing the content of the reviews (one presentation for each pair of panelists). We held a 30-minute debriefing session with each pair of panelists in which we delivered the slide presentation and asked any questions we had about the reviews. Both debriefing sessions were held by video conference.

Learning Progression Feedback: Results

We received both general and specific feedback from the panelists. As part of the general remarks, one of the panelists pointed out that no one task will address all themes or progress variables and that we should specify, for each task, which progress variables or aspects of the LP the task is intended to assess. Another panelist recommended developing task models to support the transition from the LP to task development. Task models are components of evidence-centered design (Mislevy et al., 2003), which specify the features of tasks that elicit the requisite evidence. Developing task models for an LP would entail describing the features of tasks that provide evidence that a student is at a particular level. For the long term, one of the panelists suggested that we establish links between the CoF LP and LPs for specific kinds of functions (e.g., linear functions, quadratic functions, and exponential functions). Also for the long term, we were encouraged to consider how the LP could be used as a guide for curriculum development and implementation, classroom instruction, individual student learning, diagnostic purposes, differentiated instruction, and teacher professional development.

Expert panelists proposed the following specific changes to the LP:

- The notion of function as formula is at Level 2, but evaluating expressions may occur earlier, at Level 1.
- Add clarifying language that interpreting a graph as a picture does not only occur at Level 2; rather, it is a misconception that persists from Level 2 up through Level 4—as reflected by the timeline of difficulties in Table 3.
- Keep "understand function notation as a rule" at Level 3 in the LP and move "understand the domain as a set of input values" up to Level 4 (perhaps move the high school CCSSM standards from Level 3 up to Level 4 so that Level 4 of the LP lines up with the high school CCSSM).
- Add clarifying language to Level 3 to emphasize that students at this level will not go straight from an equation to a graph without first making a table of values.
- Add clarifying language that when a function is not strictly increasing or decreasing, students at Level 4 visualize how one variable does (or does not) change with the other by describing how the graph changes over different intervals.
- Add clarifying language to indicate that slope as a rate of change is also a feature of Level 4 (and not just Level 5). Level 5 seems to indicate that any time a student refers to slope as a rate of change, the student is at Level 5. Since students are thinking about covariation at Level 4, it seems that slope as a rate of change would be a feature that students would be familiar with at Level 4.
- Update language at Level 4 to say that students can only apply procedures to compose functions and find inverses at Level 4. Level 5 seems to imply that students can only apply procedures to compose functions and find inverses at Level 4, but at Level 5, they can compose functions and find inverses with greater conceptual understanding.
- Add clarifying language to Level 5 of the LP from "students may conflate the domain and range with a particular set of inputs and outputs" to "students may conflate the domain and range of a function with a finite set of inputs and outputs that they determined."
- Function notation is not required at Grade 8 of the CCSSM (which seems to primarily correspond to Level 3). Move function notation to a higher level.

Level-Scoring Activity: Analysis and Results

Note that although the level-scoring activity was completed by the expert panelists, we refer to them here as "raters" for consistency with the rater agreement literature. The data for the level-scoring activity consisted of the probability distributions from each of 4 raters to each of 11 responses to the WW prompt "Does the graph of Wanda's walk to and from school represent a function? Explain in terms of time and her distance from home." Note that the raters provided probabilities for only 11 responses because we did not include the four responses that were used as examples in the rater packets (see Table 4). For each response and each pair of raters, we calculated two distance measures, the Kolmogorov–Smirnoff distance (Δ_{KS}) and the sum of the squared distances. Δ_{KS} is defined as follows. For rater r and response i, let $\mathbf{P}_{ri} = (P_{ri1}, P_{ri2}, \ldots, P_{rik})$ be the vector of *cumulative* probabilities, where P_{rij} is the *cumulative* probability that rater r assigned to response i for score category j. Then Δ_{KS} for raters r and s on response i is defined to be

$$\Delta_{KS}\left(\mathbf{P}_{ri}, \mathbf{P}_{si}\right) = \max\left(\left|P_{ri1} - P_{si1}\right|, \left|P_{ri2} - P_{si2}\right|, \left|P_{ri3} - P_{si3}\right|, \ldots, \left|P_{rik} - P_{sik}\right|\right). \quad (1)$$

The sum of the squared distances is defined as follows. For rater r and response i, let $\mathbf{p}_{ri} = (p_{ri1}, p_{ri2}, \ldots, p_{rik})$ be the vector of probabilities, where p_{rij} is the probability that rater r assigned to response i for score category j (note that $P_{rij} = \sum_{h=1}^{j} p_{rih}$). Then, Δ_{SS} is defined to be

$$\Delta_{SS}\left(\mathbf{p}_{ri}, \mathbf{p}_{si}\right) = \sum_{j=1}^{k}\left(p_{rij} - p_{sij}\right)^2. \quad (2)$$

For each type of distance, we calculated the mean distance across pairs of raters for each response:

$$\Delta_{KS_{ii}} = \binom{b}{2}^{-1} \sum_{r<s} \Delta\left(\mathbf{P}_{ri}, \mathbf{P}_{si}\right) \quad (3)$$

$$\Delta_{SS_{ii}} = \binom{b}{2}^{-1} \sum_{r<s} \Delta\left(\mathbf{p}_{ri}, \mathbf{p}_{si}\right), \quad (4)$$

where b is the number of raters and $1 \le r < s \le b$. In this case, there are $\binom{4}{2}$, or six pairs. The third and fourth columns of Table 7 show $\Delta_{KS_{ii}}$ and $\Delta_{SS_{ii}}$, respectively.

The responses in Table 7 are sorted in descending order of $\Delta_{KS_{ii}}$ and then in descending order of $\Delta_{SS_{ii}}$. We consider the first three responses in the table here. Response 2 split raters between Level 3 and Level 4. In their comments, the two raters who assigned a Level 3 with greater probability indicated that the response "does not specify one-valuedness" and that it "does not describe one-valuedness precisely." One of the raters who assigned Level 4 with greater probability

Table 7 Mean Distances Across All Possible Pairs of Raters for Each Response to a Prompt From Wanda's Walk

Resp.	Student response	$\Delta_{KS_{ii}}$	$\Delta_{SS_{ii}}$
2	Yes, because for every input of time, she has gone through a new output of distance from home is given.	.60	.92
4	Yes, the graph of Wanda's walk to and from school does represent a function because for every certain amount of distance, there is exactly one certain amount of time.	.53	.48
6	Yes, the graph of Wanda's walk to and from school does represent a function because the definition of a function is that every input is paired with exactly one output. And in this case, the inputs are the timer and the outputs is the distance from her house.	.48	.54
3	No, it does not. By definition of a function, every input has one output value. This is not the case between B and D where she went back closer to home and then started back to school.	.42	.53
9	Yes, Wanda's walk to and from school does represent a function because the speed in which she's traveling (input) determines how long it takes her to make the trip to and from school (output).	.35	.36
1	Yes, her distance from home is a function of time, where distance is the dependent variable and time is the independent variable. At any one time, she is only one distance from home.	.30	.24
5	Yes, because there is only one distance from home each time. Since "time" is the x-axis, and "distance from home" is the y-axis, there can be only one y-value paired to an x-value. This is a function because for each time, Wanda is only one distance from home.	.30	.24
11	Yes, because at every point in time, she is exactly and only one set distance from her house.	.30	.24
7	Yes, Wanda's walk to and from school represents a function because every input has exactly one output, meaning no input has more than one output. Therefore, by the definition of a function, Wanda's walk to and from school represents a function.	.30	.22
8	Yes, because at one certain time, she is at one distance. One input is paired with one output.	.30	.21
10	Yes, the graph is a function because it relates her distance from her home in terms of the time she spent doing this.	.22	.11

Table 8 Berry and Mielke's Kappa for Each Pair of Raters and Overall

	Probability distributions		Level with maximum probability	
Distance	Δ_{KS}	Δ_{SS}	Δ_{MU}	Δ_{MW}
Rater 1/Rater 2	.25	.40	.43	.44
Rater 1/Rater 3	.32	.49	.48	.59
Rater 1/Rater 4	.39	.56	.79	.91
Rater 2/Rater 3	.17	.37	.39	.59
Rater 2/Rater 4	.16	.30	.31	.42
Rater 3/Rater 4	.48	.51	.35	.58
Overall	.29	.43	.44	.58

noted that "likely L4 uniqueness is understood," while the other rater who assigned greater probability to Level 4 did not comment. It seems that raters were split on whether Response 2 provided evidence of one-valuedness.

Three of the four raters noted that in Response 4, the independent and dependent variables were reversed; their distributions were inclined toward lower levels than was the rater who did not make this observation. This response seems to have split raters because the error may not have been noticed by all raters, and when it was, there may have been some uncertainty as to how the error should be handled.

The disagreement for Response 6 seems to have occurred because one rater assigned a low probability to Level 4 ("The response shows weak evidence of Level 4"), while the other raters assigned a probability of .5 or greater to Level 4.

To evaluate the agreement among the multiple raters, we make use of the generalized version of Cohen's (1960) kappa developed by Berry and Mielke (1988). Let n be the number of responses. In their formulation,

$$\kappa = 1 - \frac{\delta}{\mu_\delta}, \tag{5}$$

where, for the KS distance,

$$\delta = \left[n\binom{b}{2}\right]^{-1} \sum_{i=1}^{n} \sum_{r<s} \Delta_{KS}\left(\mathbf{P}_{ri}, \mathbf{P}_{si}\right) \tag{6}$$

and

$$\mu_\delta = \left[n^2\binom{b}{2}\right]^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r<s} \Delta_{KS}\left(\mathbf{P}_{ri}, \mathbf{P}_{sj}\right). \tag{7}$$

The δ in Equation 6 can be interpreted as the observed disagreement and the μ_δ in Equation 7 as the expected disagreement. The $\Delta_{KS}(\mathbf{P}_{ri}, \mathbf{P}_{si})$s refer to the distance between distributions for corresponding responses between two raters, whereas the $\Delta_{KS}(\mathbf{P}_{ri}, \mathbf{P}_{sj})$s refer to distances between distributions from all paired combinations of responses. The definition of Berry and Mielke's kappa for Δ_{SS} can be obtained in a similar fashion with the necessary changes.

We calculated Berry and Mielke's kappas using each of Δ_{KS} and Δ_{SS} as distances, for each pair of raters and overall. The results are shown in the second and third columns of Table 8.

In addition, we calculated Berry and Mielke's kappas using two other types of distances. Let $\text{level}_{\max}(p_{ri1}, p_{ri2}, p_{ri3}, \ldots, p_{rik})$ be the score category corresponding to the maximum probability among the $\mathbf{p}_{ri}$. Since score categories can be tied for the maximum probability, let $\max(\text{level}_{\max}(p_{ri1}, p_{ri2}, p_{ri3}, \ldots, p_{rik}))$ be the highest score category among score categories with the maximum probability. Then Δ_{MU} for raters r and s on response i is defined to be

$$\Delta_{MU}\left(\mathbf{p}_{ri}, \mathbf{p}_{si}\right) = \begin{cases} 0 & \text{if } \max\left(\text{level}_{\max}\left(p_{ri1}, p_{ri2}, \ldots, p_{rik}\right)\right) = \max\left(\text{level}_{\max}\left(p_{si1}, p_{si2}, \ldots, p_{sik}\right)\right), \\ 1 & \text{otherwise,} \end{cases} \tag{8}$$

and Δ_{MW} is defined to be

$$\Delta_{MW}\left(\mathbf{p}_{ri}, \mathbf{p}_{si}\right) = \left(\max\left(\text{level}_{\max}\left(p_{ri1}, p_{ri2}, \ldots, p_{rik}\right)\right) - \max\left(\text{level}_{\max}\left(p_{si1}, p_{si2}, \ldots, p_{sik}\right)\right)\right)^2. \tag{9}$$

Kappas based on Δ_{MU} and Δ_{MW} were calculated for all pairs of raters and overall. These values are shown in the last two columns of Table 8. Note that Berry and Mielke's pairwise kappas based on Δ_{MU} are equal to Cohen's (1960) unweighted kappas between the score categories with maximum probability. Similarly, Berry and Mielke's pairwise kappas based on

Δ_{MW} are equal to quadratically weighted kappas between the score categories with maximum probability. This is because when the distance is 0 for matching scores and 1 for mismatched scores, Equation 5 is equivalent to Cohen's kappa. And when the distance is the squared difference between scores, Equation 5 is equivalent to quadratically weighted kappa.

The equivalence of these agreement statistics is shown in simplified form by Gwet (2014, par. 3.2). He shows that using the squared Euclidean distances for binary ratings of two raters that are dummy-coded using a vector in Equation 5 results in Cohen's kappa (so only one element of $\mathbf{p}_{ri}$ would be 1, and all others 0). In our case, we do not have vectors with 0 s and 1 but a vector with probabilities. However, since the definitions of Δ_{MU} and Δ_{MW} effectively reduce the probability vectors to regular rating data, the equivalence between the agreement statistics should be relatively obvious. For example, for $\mathbf{p}_{ri} = (0.1, 0.2, 0.5, 0.2)$ and $\mathbf{p}_{si} = (0.4, 0.3, 0.2, 0.1)$, the regular ratings based on the maximum probability would be 3 and 1, with $\Delta_{MU} = 1$ and $\Delta_{MW} = 4$.

Landis and Koch (1977) proposed a set of benchmarks for evaluating kappa. Less than 0.00 is considered poor agreement, 0.00 – 0.20 is slight, 0.21 – 0.40 is fair, 0.41 – .0.60 is moderate, 0.61 – 0.80 is substantial, and 0.81 – 1.00 is almost perfect agreement. The kappas based on Δ_{SS} are higher than the kappas based on Δ_{KS}, and the kappas based on Δ_{MW} are higher than the kappas based on Δ_{MU}, suggesting that when raters disagree, the disparity in their disagreements tends to be smaller rather than larger. With one exception, the kappas based on Δ_{MU} are higher than the kappas based on Δ_{KS}, and the kappas based on Δ_{MW} are always higher than the kappas based on Δ_{SS}, suggesting that agreement is higher on the scores corresponding to the maximum probabilities than on how probabilities are distributed across scores. The kappas with Rater 2 tend to be lower than the kappas with the other raters.

Summary

We evaluated the CoF LP using three sets of results: student responses to cognitive interviews, recommendations from an expert panel, and probability distributions from a level-scoring activity, also provided by the panel. We now address each of the research questions that we outlined in the introduction.

Research Question 1

Task performance on the cognitive interviews was in general very high. This was somewhat surprising, given the findings that misconceptions about CoF continue through the college years. However, this sample consisted of highly motivated, high-achieving students. Performance was likely further bolstered by access to the definitions of function. This was done because we wanted to assess student understanding of the CoF rather than the ability to recall a definition.

We did not see very many instances of performance typical of earlier levels of the LP. Nevertheless, some misconceptions were evident. One of the main findings from the cognitive interviews is that at least through Level 4, students continue to connect points with line segments when a function is only defined at certain points. Although this misconception is documented in the literature (Leinhardt et al., 1990), one of the panelists commented that it is rarely addressed by the curriculum, which is a possible explanation for why it persists into the higher levels. As a result, we modified the LP to make it clear that this misconception persists through Level 4.

Another finding from the cognitive interviews is that a few students rejected functions as functions on the grounds that they were many-to-one. This misconception is addressed by the LP and persists through Level 3. We also saw evidence of rejecting a function as such because "it has no rule" (see the second example in Table 5). This is also addressed at Level 3 of the LP. Also, there seemed to be some consistency on the part of students who rejected functions as functions: three of the four students who rejected the function in WW as such also at some point rejected the function in AB as such. Three of four students who responded incorrectly to the questions of whether the encryption rule and decryption rule is each a function in SM also indicated that at least one of the representations in WW and/or AB is not a function. These findings would tend to support the specification of levels in the LP.

On a more task-specific note, there were a few instances where students did not answer both parts of Questions 3, 4, and 5 of task SM. This can be addressed by breaking each of these questions into two parts, for example, Question 3 can be split into two questions, as follows:

3a. Is the mapping from the plaintext to the ciphertext a function? Explain.
3b. Is the mapping from the ciphertext to the plaintext a function? Explain.

Table 9 Mappings Between Tasks Wanda's Walk, Annika's Bakery, and Secret Messages and Progress Variables

Task	Progress variable assessed	Progress variable demand
Wanda's Walk		
a	Covariation/interpretation	Pointwise/static
b	Multiple representations	Integrated (graph → text description)
	Covariation/interpretation	Global/dynamic
c	Covariation/interpretation	Global/dynamic
d	Function definition/domain & range	As outputs dependent on inputs
Annika's Bakery		
a	Function definition/domain & range	As outputs dependent on inputs
b	Multiple representations	Integrated (text description → table)
	Covariation/interpretation	Pointwise/static
c	Function definition/domain & range	As outputs dependent on inputs
d	Multiple representations	Integrated (table → graph)
	Covariation/interpretation	Pointwise/static
e	Function definition/domain & range	As outputs dependent on inputs
f	Multiple representations	Integrated (text description/table/graph → arrow diagram)
Secret Messages		
1, 2	Covariation/interpretation	Pointwise/static
3, 4, 5	Function definition/domain & range	As outputs dependent on inputs

Research Question 2

The expert panel was positive about the LP, however, they did give both general and specific suggestions for revision. We made all of the specific suggestions recommended by the panel, except for the suggestion to "add clarifying language to Level 5 of the LP from 'students may conflate the domain and range with a particular set of inputs and outputs' to 'students may conflate the domain and range of a function with a finite set of inputs and outputs that they determined.'" This misconception was based on expert opinion rather than findings from the literature, and on further discussion, we decided there was insufficient evidence for its existence, so we removed it from the LP.

Most of the general suggestions pertain to long-term revisions or extensions to the existing work (e.g., the development of task models). As yet we have not made these revisions. However, we did address the suggestion to make explicit the themes or progress variables from the LP that are assessed by the existing tasks. These mappings are shown in Table 9. The first column shows the task (tasks are subdivided into parts), the second column shows the progress variable(s) assessed by the task part (progress variables are given in Figure 1), and the last column shows the level of the progress variable demanded by the task (also from Figure 1). For example, identifying at which points Wanda is at home requires only pointwise/static interpretation, but describing what she does across time intervals requires global/dynamic interpretation. Apart from recommending the development of task models and mapping the tasks to progress variables, experts did not suggest any revisions to the tasks.

Research Question 3

Some of the student responses from the cognitive interviews seemed straightforward to classify using the levels of the LP, while other responses seemed less canonical, showing features of multiple levels of the LP. We believe this is to a large extent unavoidable, since as has been noted by others, an LP does not assume that students always perform consistently at a particular level. It was for this reason that we asked panelists to assign probability distributions, rather than single scores, to each student response.

Research Question 3a

The results from the level-scoring activity point to a few responses that were not classified consistently among raters, as reflected by the average distances between pairs of raters' probability distributions. For one of the responses, raters were split as to whether the student showed understanding of one-valuedness. Another response was not consistently classified because raters had a difference of opinion on how to treat the reversal of the independent and dependent variables. While

Table 10 Revised Version of Timeline of Concepts and Skills

Skill/concept	Level 1	Level 2	Level 3	Level 4
Extend patterns in one dimension	X	X	X	X
Evaluate formulas	X	X	X	X
Plot points on a graph		X	X	X
Detect increasing/decreasing trends		X	X	X
Interpret function notation			X	X
Understand notion of dependency			X	X
Recognize equivalent representations			X	X
Understand domain as a set of inputs				X
Understand range as a set of outputs				X
Understand covariational reasoning				X
Understand "one-valuedness"				X
Compose and find inverses				X

these results suggest something about the nature of responses for which there may not be consensus, we do not think that by themselves they implicate revision to the LP. If we see evidence that many students are consistently making the same kinds of errors (e.g., reversing the independent and dependent variables) when we evaluate the LP with a greater number of tasks and a larger sample size, then we may modify the LP to account for this.

Research Question 3b

With respect to the level-scoring activity, agreement among raters is higher when the kappas are based on distances between scores with the maximum probabilities than when they are based on distances between classification probabilities. It has been suggested in the literature that student performances may show aspects of multiple levels, but raters do not agree very well on what the classification probabilities across levels should be. Rather, they agree better on the most likely level of performance.

The revised LP, based on the results of the cognitive interviews and the feedback from the panel, is shown in Appendix A. Some of these changes also required modifications to the timeline of concepts and skills, the revised version of which is shown in Table 10.

The results from the cognitive interviews and recommendations from the expert panel informed initial revisions to the LP. A next step is to develop task models for additional tasks that assess both higher and lower levels of the LP. It may also be advisable to conduct additional cognitive interviews with a diverse sample of students. This would be followed by a large-scale empirical recovery study. If the levels of the LP are empirically recovered, we will follow the recommendation of one of our panelists and propose how the LP could be used as a guide for curriculum development and implementation, classroom instruction, individual student learning, diagnostic purposes, differentiated instruction, and teacher professional development. This would pave the way for an evaluation of instructional efficacy, the last step in the cycle for validating an LP. Such as evaluation could be carried out through teaching experiments, as proposed by Shavelson (2009). Such an experiment might involve comparing LP-based instruction to traditional instruction.

Acknowledgments

Thanks to Dan McCaffrey for providing suggestions for the analysis, to Sarah Ohls for help in conducting cognitive interviews, and to the expert panelists for their suggestions. We are grateful for their contributions, and any errors are the sole responsibility of the authors.

Notes

1 CCSSM, high school Standard F-IF.5.
2 CCSSM, high school Standard F-IF.2.
3 CCSSM, high school Standard F-IF.4, brackets added.
4 CCSSM, high school Standard F-IF.6.

5 CCSSM, high school Standard F-IF.1.

6 CCSSM, high school Standard F-IF.3. Note that understanding sequences as examples of functions is very different from working with sequences. It is expected that students would work with sequences early on (they are usually students' first exposure to functions). Students can extend sequences at Level 1.

7 CCSSM, High School Standard F-IF.7.

8 CCSSM, p. 67.

9 CCSSM, eighth-grade Standard 8.F.2.

10 CCSSM, eighth-grade Standard 8.F.5.

11 CCSSM, eighth-grade Standard 8.F.1.

12 See Carlson and Oehrtmann (2005).

13 According to Leinhardt et al. (1990), functions of time are both more common and easier to grasp.

14 See Vergnaud (1983) for the distinction between scalar and functional strategies.

15 CCSSM, high school Standard F-IF.5.

16 CCSSM, high school Standard F-IF.2.

17 CCSSM, high school Standard F-IF.4, brackets added.

18 CCSSM, high school Standard F-IF.7.

19 CCSSM, high school Standard F-IF.6.

20 CCSSM, high school Standard F-IF.1.

21 CCSSM, high school Standard F-IF.3. Note that understanding sequences as examples of functions is very different from working with sequences. It is expected that students would work with sequences early on (they are usually students' first exposure to functions). Students can extend sequences at Level 1.

22 See Carlson and Oehrtmann (2005)

23 CCSSM, p. 67.

24 CCSSM, eighth-grade Standard 8.F.2.

25 CCSSM, eighth-grade Standard 8.F.5.

26 CCSSM, eighth-grade Standard 8.F.1.

27 According to Leinhardt et al. (1990), functions of time are both more common and easier to grasp.

28 See Vergnaud (1983) for the distinction between scalar and functional strategies.

References

Arieli-Attali, M., & Cayton-Hodges, G. (2014). *Expanding the CBAL mathematics assessments to elementary grades: The development of a competency model and a rational number learning progression* (Research Report No. RR-14-08). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12008

Arieli-Attali, M., Wylie, E. C., & Bauer, M. I. (2012, April). *The use of three learning progressions in supporting formative assessment in middle school mathematics*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, BC.

Battista, M. T. (2011). Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *The Mathematics Enthusiastic, 8*, 507–570.

Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning: A preliminary theory of action for summative and formative assessment. *Measurement: Interdisciplinary Research and Perspectives, 8*, 70–91. https://doi.org/10.1080/15366367.2010.508686

Bennett, R. E., & Gitomer, D. H. (2009). Transforming K–12 assessment: Integrating accountability testing, formative assessment, and professional support. In C. Wyatt-Smith & J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 43–61). New York, NY: Springer. https://doi.org/10.1007/978-1-4020-9964-9_3

Berry, K. J., & Mielke, P. W. (1988). A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters. *Educational and Psychological Measurement, 48*, 921–933. https://doi.org/10.1177/0013164488484007

Carlson, M., & Oehrtman, M. (2005). *Research sampler 9: Key aspects of knowing and learning the concept of function*. Retrieved from http://www.maa.org/programs/faculty-and-departments/curriculum-department-guidelines-recommendations/teaching-and-learning/9-key-aspects-of-knowing-and-learning-the-concept-of-function

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, *6*, 81–89. https://doi.org/10.1207/s15327833mtl0602_1

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46. https://doi.org/10.1177/001316446002000104

Collins, W., Cuevas, G., Foster, A. G., Gordon, B., Moore-Harris, B., Rath, J., . . . Winters, L. J.. (2001). *Glencoe algebra 1: Integration, applications, connections*. New York, NY: Glencoe/McGraw-Hill.

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics (CCSSM)*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. Retrieved from http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf

Confrey, J., Maloney, A., Nguyen, K., Mojica, G., & Myers, M. (2009). *Equipartitioning/splitting as a foundation of rational number reasoning using learning trajectories*. Paper presented at the 33rd conference of the International Group for the Psychology of Mathematics Education, Thessaloniki, Greece.

Confrey, J., & Smith, E. (1994). Exponential functions, rates of change, and the multiplicative unit. *Educational Studies in Mathematics*, *26*, 135–164. https://doi.org/10.2307/749228

Confrey, J., & Smith, E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education*, *26*, 66–86. https://doi.org/10.2307/749228

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence-based approach to reform* (Research Report No. RR-63). Philadelphia, PA: Consortium for Policy Research in Education.

Daro, P., Mosher, F. A., & Corcoran, T. (2011). *Learning trajectories in mathematics: A foundation for standards, curriculum, assessment, and instruction* (Research Report No. RR-68). Philadelphia: Consortium for Policy Research in Education.

Dubinsky, E., & Harel, G. (1992). The nature of the process conception of function. In *The Concept of function: Aspects of epistemology and pedagogy* (MAA Notes No. 25, pp. 85–106). Washington, DC: Mathematical Association of America.

Dubinsky, E., & Wilson, R. T. (2013). High school students' understanding of the function concept. *The Journal of Mathematical Behavior, 32*, 83–101. https://doi.org/10.1016/j.jmathb.2012.12.001

Dugdale, S. (1993). Functions and graphs: Perspectives on student thinking. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representation of functions* (pp. 101–130). Hillsdale, NJ: Erlbaum.

Educational Testing Service. (n.d.). *The CBAL mathematics competency model and provisional learning progressions*. Retrieved from https://www.ets.org/cbal/mathematics/

Graf, E. A. (2009). *Defining mathematics competency in the service of cognitively based assessment for Grades 6 through 8* (Research Report No. RR-09-42). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2009.tb02199.x

Graf, E. A. (2015a). *Annika's bakery* [Mathematics task]. Princeton, NJ: Educational Testing Service.

Graf, E. A. (2015b). *Secret messages* [Mathematics task]. Princeton, NJ: Educational Testing Service.

Graf, E. A., & Arieli-Attali, M. (2015). Designing and developing assessments of complex thinking in mathematics for the middle grades. *Theory Into Practice, 54*, 195–202. https://doi.org/10.1080/00405841.2015.1044365

Graf, E. A., Harris, K., Marquez, E., Fife, J., & Redman, M. (2009). *Cognitively Based Assessment of, for, and as Learning (CBAL) in mathematics: A design and first steps toward implementation* (Research Memorandum No. RM-09-07). Princeton, NJ: Educational Testing Service.

Graf, E. A., Harris, K., Marquez, E., Fife, J., & Redman, M. (2010). Highlights from the Cognitively Based Assessment of, for, and as Learning (CBAL) project in mathematics. *ETS Research Spotlight, 3*, 19–30.

Graf, E. A., & van Rijn, P. W. (2016). Learning progressions as a guide for design: Recommendations based on observations from a mathematics assessment. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.), *Handbook of test development* (2nd ed., pp. 165–189). New York, NY: Routledge.

Gwet, K. L. (2014). *Handbook of inter-rater reliability* (4th ed.). Gaithersburg, MD: Advanced Analytics LLC.

Kalchman, M., & Koedinger, K. (2005). Teaching and learning functions. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History, mathematics and science in the classroom* (pp. 351–396). Washington, DC: National Academies Press.

Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for development of mathematical understanding: Rational numbers and functions. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 1–38). Mahwah, NJ: Erlbaum.

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (pp. 17–64). Westport, CT: Praeger.

Kane, M. T. (2013). Validation as a pragmatic, scientific activity. *Journal of Educational Measurement, 50*, 115–122. https://doi.org/10.1111/jedm.12007

Kane, M., & Bejar, I. I. (2014). Cognitive frameworks for assessment, teaching, and learning: A validity perspective. *Psicología Educativa, 20*, 117–123. https://doi.org/10.1016/j.pse.2014.11.006

Karplus, R. (1979). Continuous functions: Students' viewpoints. *European Journal of Science Education, 1*, 397–415. https://doi.org/10.1080/0140528790010404

Kieran, C. (1993). Functions, graphing, and technology: Integrating research on learning and instruction. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representation of functions* (pp. 189–237). Hillsdale, NJ: Erlbaum.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*, 1–64. https://doi.org/10.3102/00346543060001001

Markovits, Z., Eylon, B. S., & Bruckheimer, M. (1986). Functions today and yesterday. *For the Learning of Mathematics, 6*(2), 18–28.

Marquez, E. (2015). *Wanda's Walk* [Mathematics task]. Princeton, NJ: Educational Testing Service.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives, 1*, 3–67. https://doi.org/10.1207/S15366359MEA0101_02

Monk, S. (1992). Students' understanding of a function given by a physical model. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (MA Nores No. 25, pp. 175–194). Washington, DC: Mathematical Association of America.

National Assessment Governing Board. (2008). *Science framework for the 2009 National Assessment of Educational Progress*. Washington, DC: Author.

Plake, B. S., Huff, K., & Reshetar, R. (2010). Evidence-centered assessment design as a foundation for achievement-level descriptor development and for standard setting. *Applied Measurement in Education, 23*, 342–357. https://doi.org/10.1080/08957347.2010.510964

Ponce, G. A. (2007). Critical juncture ahead! Proceed with caution to introduce the concept of function. *Mathematics Teacher, 101*, 136–144.

Riconscente, M. M., Mislevy, R. J., & Corrigan, S. (2016). Evidence-centered design. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.), *Handbook of test development* (2nd ed., pp. 40–63). New York, NY: Routledge.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1–36. https://doi.org/10.1007/BF00302715

Sfard, A. (1992). Operational origins of mathematical notions and the quandary of reification: The case of function. In E. Dubinsky & G. Harel (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (MAA Notes No. 25, pp. 59–84). Washington, DC: Mathematical Association of America.

Shavelson, R. J. (2009, June). *Reflections on learning progressions*. Paper presented at the Learning Progressions in Science (LeaPS) conference, Iowa City, IA.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*, 114–145. https://doi.org/10.2307/749205

Smith, C., Wiser, M., Anderson, C. W., Krajcik, J., & Coppola, B. (2004). *Implications of research on children's learning for assessment: Matter and atomic molecular theory*. Washington, DC: Center for Education, National Research Council.

van Rijn, P. W., Graf, E. A., Arieli-Attali, M., & Song, Y. (2018). *Agreement of teachers on evaluating assessments of learning progressions in English language arts and mathematics* (Research Report No. RR-18-11). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12199

van Rijn, P. W., Graf, E. A., & Deane, P. (2014). Empirical recovery of argumentation learning progressions in scenario-based assessments of English language arts. *Psicología Educativa, 20*, 109–115. https://doi.org/10.1016/j.pse.2014.11.004

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 127–174). New York, NY: Academic Press.

Vinner, S., & Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for Research in Mathematics Education, 20*, 356–366. https://doi.org/10.2307/749441

Wilmot, D. B., Schoenfeld, A., Wilson, M., Champney, D., & Zahner, W. (2011). Validating a learning progression in mathematical functions for college readiness. *Mathematical Thinking and Learning, 13*, 259–291. https://doi.org/10.1080/10986065.2011.608344

Appendix A: Concept of Function Learning Progression

Concept of Function Learning Progression Prior to Review

Level 6: Drawing Extensions

Full Abstraction

A function is perceived as an object that can be operated upon; understanding of domain and range is also well-developed

At this level students can recognize, produce, and work with nondeterministic functions and functions with discontinuities or exceptional points like corners.

At this level students may have slips, but there wouldn't be any consistent difficulties.

This level represents a professional mathematician's level of understanding of the concept of *function*. Students at this level understand the formal set-theoretic definition of function and connect it to core ideas about inputs and outputs and covariational reasoning that have been developed earlier. At this level students recognize families of functions as parameterized objects and understand the role of the parameters within those families. Students at this level take greater care with the ideas of domain and range, understanding that changing the domain of a function changes the function. They can "Relate the domain of a function to its graph and, where applicable, to the quantitative relationship it describes. For example, if the function $h(n)$ gives the number of person-hours it takes to assemble n engines in a factory, then the positive integers would be an appropriate domain for the function.★"[1]

Students at this level are able to view functions as objects or points in a space. Students understand the algebra of functions. For example, if f and g are real-valued functions, then $f+g$ and $f{\cdot}g$ are defined as $(f+g)(x)=f(x)+g(x)$ and $(f{\cdot}g)(x)=f(x){\cdot}g(x)$; alternatively, if f and g are functions from a set to itself, then $f{\cdot}g$ can be defined by $(f{\cdot}g)(x)=f(g(x))$.

Level 5: Thorough Conceptualization

Object Perception of a Function

A function is perceived as an object that can be operated upon

At this level students can comfortably use function notation, and can transform functions. As expressed in the CCSSM: "[Students can] Use function notation, evaluate functions for inputs in their domains, and interpret statements that use function notation in terms of a context."[2] Since students attend to global features of graphs, they are unlikely to make any of the more common graphing errors that emerge at Level 2. "For a function that models a relationship between two quantities, [they can] interpret key features of graphs and tables in terms of the quantities, and sketch graphs showing key features given a verbal description of the relationship. Key features include: intercepts; intervals where the function is increasing, decreasing, positive, or negative; relative maximums and minimums; symmetries; end behavior; and periodicity.★"[3] They are just starting to understand the role of parameters in some function families to the extent that they can recognize how changes in some of these parameters influence the behavior of graphs. In addition, they are unlikely to overgeneralize linear or quadratic functions, but can apply linear approximation when appropriate. Students can: "Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval, [and] estimate the rate of change from a graph.★"[4] Students compose functions and find inverses not just by applying procedures (such as switching x and y and solving for y or reflecting a graph across the line $y=x$), but with greater conceptual understanding.

Continued

At this level students have difficulty with Although students at this level understand that a function is not necessarily given by a formula, they may still have difficulty recognizing some kinds of functions (e.g., nondeterministic functions or discontinuous functions). Also characterizing the domain and range may also still pose difficulty, for example, students may conflate the domain and range with a particular set of inputs and outputs.

At this level, students think of functions as expressed in the CCSSM: "Understand that a function from one set (called the domain) to another set (called the range) assigns to each element of the domain exactly one element of the range. If f is a function and x is an element of its domain, then $f(x)$ denotes the output of f corresponding to the input x. The graph of f is the graph of the equation $y = f(x)$."[5] ***Rather than abandoning what they have learned about the notion of dependence, students at this level build on it, and apply the CCSSM definition of function flexibly, as the situation demands***. An additional important understanding that is acquired at this level is the ability to identify sequences as functions. This is expressed in the CCSSM standard: "Recognize that sequences are functions, sometimes defined recursively, whose domain is a subset of the integers. For example, the Fibonacci sequence is defined recursively by $f(0) = f(1) = 1, f(n+1) = f(n) + f(\mathrm{n}-1)$ *for* $n \geq 1$."[6]

Since they now conceive of a ***function as an object***, students are beginning to understand the role of parameters in some families of functions (e.g., they understand how changing the slope or the intercept of a linear function changes its graph).

Level 4: Synthesis

Full Covariational Relationship
Scheme of a function as a dependency relationship between quantities, with full linking between alternate representations

At this level students can skillfully use alternative equivalent representations, and can identify features that are common to equivalent representations (e.g., they can identify a y-intercept by looking at a graph, equation, or table). They can "Graph functions expressed symbolically and show key features of the graph, by hand in simple cases and using technology for more complicated cases.★"[7] Because they are thinking in terms of inputs and outputs, students at this level can start to ***compose functions*** and ***find their inverses***. Students at this level are working successfully with functions for which the independent variable is not necessarily time (or dependent on time), and they can start to find the equation from a table or a graph for simpler functions.

At this level students have difficulty with still overgeneralizing linear functions but only in more subtle situations (e.g., interpolation between points), and linear approximation is applied as a technique rather than as an overgeneralization. They may still have difficulty expressing rules in formal notation.

At this level, students have a fully developed notion of dependence; that is that a function operates on inputs to yield outputs. Their interpretation of functions is no longer pointwise; they think about how the dependent variable changes with the independent variable and the specific nature of that change; in other words, ***they think about how variables covary***. Interpretation of domain and range has not changed from Level 3. Students accept that a function may be associated with more than one computational process (though it is not a requirement that they be able to prove their equivalence). Students at this level are ***attending to the meaning of features*** of graphs of functions such as minima, maxima, and asymptotes. Students at this level apply the "one-valuedness" core idea in identifying functions (that is, that each input is assigned to one and only one output). In other words, they have a complete understanding of the CCSSM 8.F definition of function (see Level 3). Students at this level are using patterns of change to select an appropriate function to model a situation; i.e., they can start to find the equation from a table for simpler functions. At this level students fully

Continued

recognize that 'a function can be described in various ways, such as by a graph (e.g., the trace of a seismograph); by a verbal rule, as in, "I'll give you a state, you give me the capital city;" by an algebraic expression like $f(x) = a + bx$; or by a recursive rule.'[8] At this level they are fluent with "Compar[ing] properties of two functions each represented in a different way (algebraically, graphically, numerically in tables, or by verbal descriptions)."[9] Students at this level can: "Describe qualitatively the functional relationship between two quantities by analyzing a graph (e.g., where the function is increasing or decreasing, linear or nonlinear). [They can also] sketch a graph that exhibits the qualitative features of a function that has been described verbally."[10]

Level 3: Making Connections

Generalized Relational-Function as a Rule
Scheme of a relationship between quantities <u>*rule-wise*</u>, bound to specific representation, with early linking between some alternate representations

At this level students can start to see the link between different representations of a function (equation, table, graph, verbal description), and recognize that different, but equivalent formulas can be used to represent the same function, even if they cannot always demonstrate their equivalence. They also recognize that rules for functions are not necessarily characterized by formulas. They can more easily translate from equations to graphs and tables than the reverse.

At this level students have difficulty with generating an equation from a table or a graph, i.e., producing the reverse translation between representations, even though they may be able to recognize it. When graphing, students at Level 3 may make some of the same errors as at Level 2, though less frequently. At this level, students might still reject the following kinds of functions as such: (1) Many-to-one functions, (2) piecewise functions, (3) discontinuous functions, (4) functions with exceptional points, and (5) nondeterministic functions. Students at this level still overgeneralize linear functions to nonlinear situations.

At this level, students are starting to think of functions as expressed in the CCSSM, though they have not yet developed the one-valuedness aspect of the definition: "Understand that a function is a rule that assigns to each input exactly one output. The graph of a function is the set of ordered pairs consisting of an input and the corresponding output."[11] They have an operational notion of function, and are able to apply consistent ***rules*** to find outputs given inputs. In other words, a notion of ***dependence*** and outputs generated from inputs starts to develop at Level 3. They understand the domain as a set of input values (or set of first coordinates in an ordered pair), and the range as a set of output values (or the set of second coordinates in an ordered pair).

Students at this level are more easily working with functions of ***time***, since these kinds of functions follow a chronological sequence that is more familiar from everyday life. Because they are starting to think in terms of inputs and outputs, function notation is less problematic, though some errors are still common without scaffolding support (e.g., $f(g(5))$ may be interpreted as "the value of f when g is 5," and f and g may be confused with variables.[12] Though their interpretation of functions is still primarily pointwise, they are starting to distinguish among types of functions by observing the pattern of change.

Level 2: Familiarization

Local Relational—Function as a Formula
Scheme of a relationship between quantities pointwise, bound to a computational process or algebraic expression

At this level students can use an equation to generate ordered pairs in a table and/or plot points on a graph, but this is carried out as a sequence of procedures. However, students can grasp the trend in simple cases, namely, they can recognize when a dependent variable changes with an independent variable (the mutual change), and may be able to characterize it as strictly increasing or strictly decreasing.

Continued

At this level students have difficulty with perceiving the connections between the different representations of a function. Because graphing is pointwise, common errors that may be in evidence include (1) seeing only marked points as "part" of a graph; (2) connecting points when a graph is not continuous; (3) neglecting considerations of scale; (4) when interpolating or extrapolating points on a graph, reverting to simple examples, and overgeneralizing linearity; and (5) confusing relationships with pictorial aspects of the situation.

At this early stage of learning about functions, students are developing a ***scheme*** for the relationship between two sets of numbers, but it is pointwise and based on operations. Students at this level think of a function as an equation or a computational process, and do not perceive a link between alternate representations of the same function, that is, they can work with the equation, the table, or the graph of a function, but not necessarily understand the connections between those different representations. Even equivalent functions (equations) are not necessarily perceived as such when alternate variable names are used, and alternate computational processes (e.g., a function expressed in explicit form or recursive form) are definitely not perceived as the same function. They can evaluate expressions, but do not yet understand formal function notation (e.g., what f and $f(x)$ mean), nor do they conceive of a function as mapping from inputs to outputs.[13] It is important to note that while a formula is one kind of rule for a function, it is certainly not the only kind. At this level, students might still reject a function defined by a graph as such. It is not until Level 3 that students recognize that the rule for a function is not necessarily a formula.

Level 1: Preexposure

Quantity Perception
Perceiving a quantity as is, and operating on it

At this level students can work with sequences by extending them or finding missing terms. They can complete patterns. Although they would not use function notation, they can informally express recursive rules for sequences (e.g., "Keep adding 3"). They would not express rules in explicit form (take the term as input and generate the corresponding value as output).

At this level students have difficulty with thinking relationally, namely connecting a pattern in one dimension to a pattern in another.

Students at this level have a solid perception and understanding of patterns in one dimension, attending to quantities and operating on them. At this stage, students had not yet been introduced to functions, nor do they perceive a relation or mapping between two patterns. In the context of proportional reasoning, students may employ scalar strategies but not functional strategies.[14]

Revised Version of the Concept of Function Learning Progression

Level 6: Drawing Extensions	
At this level students can recognize, produce, and work with nondeterministic functions and functions with discontinuities or exceptional points like corners. **At this level students may have slips**, but there wouldn't be any consistent difficulties. This level represents a professional mathematician's level of understanding of the concept of function. Students at this level understand the formal set-theoretic definition of function and connect it to core ideas about inputs and outputs and covariational reasoning that have been developed earlier. At this level students recognize families of functions as parameterized objects and understand the role of the parameters within those families. Students at this level take greater care with the ideas of domain and range, understanding that changing the domain of a function changes the function. They can "Relate the domain of a function to its graph and, where applicable, to the quantitative relationship it describes. For example, if the function $h(n)$ gives the number of person-hours it takes to assemble n engines in a factory, then the positive integers would be an appropriate domain for the function.★"[15] Students at this level are able to view functions as objects or points in a space. Students understand the algebra of functions. For example, if f and g are real-valued functions, then $f+g$ and $f \cdot g$ are defined as $(f+g)(x)=f(x)+g(x)$ and $(f \cdot g)(x)=f(x) \cdot g(x)$; alternatively, if f and g are functions from a set to itself, then $f \cdot g$ can be defined by $(f \cdot g)(x)=f(g(x))$.	***Full Abstraction*** A function is perceived as an object that can be operated upon; understanding of domain and range is also well-developed
Level 5: Thorough Conceptualization	
At this level students can comfortably use function notation, and can transform functions. As expressed in the CCSSM: "[Students can] Use function notation, evaluate functions for inputs in their domains, and interpret statements that use function notation in terms of a context."[16] Students attend to global features of graphs, and are unlikely to make any of the more common graphing errors that emerge at Level 2, such as connecting points when a graph is not continuous. "For a function that models a relationship between two quantities, [they can] interpret key features of graphs and tables in terms of the quantities, and sketch graphs showing key features given a verbal description of the relationship. Key features include: intercepts; intervals where the function is increasing, decreasing, positive, or negative; relative maximums and minimums; symmetries; end behavior; and periodicity.★"[17] Consistent with this, students can also "Graph functions expressed symbolically and show key features of the graph, by hand in simple cases and using technology for more complicated cases.★"[18] They are just starting to understand the role of parameters in some function families to the extent that they can recognize how changes in some of these parameters influence the behavior of graphs. In addition, they are unlikely to overgeneralize linear or quadratic functions, but can apply linear approximation when appropriate. Students can: "Calculate and interpret the average rate of change of a function (presented symbolically or as a table) over a specified interval, [and] estimate the rate of change from a graph.★"[19] Students compose functions and find inverses not just by applying procedures (such as switching x and y and solving for y or reflecting a graph across the line $y=x$), but with greater conceptual understanding. Students at this level would no longer confuse graphs with pictorial aspects of the situation.	***Object Perception of a Function*** A function is perceived as an object that can be operated upon

Continued

At this level students have difficulty with Although students at this level understand that a function is not necessarily given by a formula, they may still have difficulty recognizing some kinds of functions (e.g., nondeterministic functions or discontinuous functions). Also characterizing the domain and range may still pose some difficulty.

At this level, students think of functions as expressed in the CCSSM: "Understand that a function from one set (called the domain) to another set (called the range) assigns to each element of the domain exactly one element of the range. If f is a function and x is an element of its domain, then $f(x)$ denotes the output of f corresponding to the input x. The graph of f is the graph of the equation $y = f(x)$."[20] ***Rather than abandoning what they have learned about the notion of dependence, students at this level build on it, and apply the CCSSM definition of function flexibly, as the situation demands***. An additional important understanding that is acquired at this level is the ability to identify sequences as functions. This is expressed in the CCSSM standard: "Recognize that sequences are functions, sometimes defined recursively, whose domain is a subset of the integers. For example, the Fibonacci sequence is defined recursively by $f(0) = f(1) = 1, f(n + 1) = f(n) + f(n - 1)$ for $n \geq 1$."[21]

Since they now conceive of a ***function as an object***, students are beginning to understand the role of parameters in some families of functions (e.g., they understand how changing the slope or the intercept of a linear function changes its graph).

Level 4: Synthesis

Full Covariational Relationship

Scheme of a function as a dependency relationship between quantities, with full linking between alternate representations

At this level students can skillfully use alternative equivalent representations, and can identify features that are common to equivalent representations (e.g., they can identify a y-intercept by looking at a graph, equation, or table). Because they are thinking in terms of inputs and outputs, students at this level can start to ***compose functions*** and ***find their inverses***, though they do so primarily by relying on memorized procedures and without conceptual understanding. They understand slope as a rate of change. Students at this level are working successfully with functions for which the independent variable is not necessarily time (or dependent on time), and they can start to find the equation from a table or a graph for simpler functions.

At this level students have difficulty with still overgeneralizing linear functions but only in more subtle situations (e.g., interpolation between points), and linear approximation is applied as a technique rather than as an overgeneralization. They can still confuse graphs with pictorial aspects of the situation. Although they have been introduced to formal function notation at this level, students may still have difficulty using it to express rules. For example, $f(g(5))$ may be interpreted as "the value of f when g is 5," and f and g may be confused with variables.[22]

At this level, students have a fully developed notion of dependence; that is that a function operates on inputs to yield outputs. Their interpretation of functions is no longer pointwise; they think about how the dependent variable changes with the independent variable and the specific nature of that change; in other words, ***they think about how variables covary***, and understand slope as a rate of change. They understand the domain as a set of input values (or set of first coordinates in an ordered pair), and the range as a set of output values (or the set of second coordinates in an ordered pair). Still, they have difficulty fully characterizing the domain and range. Students accept that a function may be associated with more than one computational process (though it is not a requirement that they be able to prove their equivalence). Students at this level are ***attending to the meaning of features*** of

Continued

graphs of functions such as minima, maxima, and asymptotes. Students at this level apply the "one-valuedness" core idea in identifying functions (that is, that each input is assigned to one and only one output). In other words, they have a complete understanding of the CCSSM 8.F definition of function (see Level 3). Students at this level are using patterns of change to select an appropriate function to model a situation; i.e., they can start to find the equation from a table for simpler functions. At this level students fully recognize that 'a function can be described in various ways, such as by a graph (e.g., the trace of a seismograph); by a verbal rule, as in, "I'll give you a state, you give me the capital city;" by an algebraic expression like $f(x) = a + bx$; or by a recursive rule.'[23] At this level they are fluent with "Compar[ing] properties of two functions each represented in a different way (algebraically, graphically, numerically in tables, or by verbal descriptions)."[24] Students at this level can: "Describe qualitatively the functional relationship between two quantities by analyzing a graph (e.g., where the function is increasing or decreasing, linear or nonlinear). [They can also] sketch a graph that exhibits the qualitative features of a function that has been described verbally."[25]

Level 3: Making Connections

Generalized Relational-Function as a Rule
Scheme of a relationship between quantities rule-wise, bound to specific representation, with early linking between some alternate representations

At this level students can start to see the link between different representations of a function (equation, table, graph, verbal description), and recognize that different, but equivalent formulas can be used to represent the same function, even if they cannot always demonstrate their equivalence. They also recognize that rules for functions are not necessarily characterized by formulas. They can more easily translate from equations to graphs and tables than the reverse. When generating graphs from equations, students at this level would first generate an input–output table.

At this level students have difficulty with generating an equation from a table or a graph, i.e., producing the reverse translation between representations, even though they may be able to recognize it. When graphing, students at Level 3 may make the same errors as at Level 2, though less frequently. At this level, students might still reject the following kinds of functions as such: (1) Many-to-one functions, (2) piecewise functions, (3) discontinuous functions, (4) functions with exceptional points, and (5) nondeterministic functions. Students at this level still overgeneralize linear functions to nonlinear situations, and they can still confuse graphs with pictorial aspects of the situation. Students at this level may not have been introduced to formal function notation.

At this level, students are starting to think of functions as expressed in the CCSSM, though they have not yet developed the one-valuedness aspect of the definition: "Understand that a function is a rule that assigns to each input exactly one output. The graph of a function is the set of ordered pairs consisting of an input and the corresponding output."[26] They have an operational notion of function, and are able to apply consistent ***rules*** to find outputs given inputs. In other words, a notion of ***dependence*** and outputs generated from inputs starts to develop at Level 3.

Students at this level are more easily working with functions of ***time***, since these kinds of functions follow a chronological sequence that is more familiar from everyday life. Though their interpretation of functions is still primarily pointwise, they are starting to distinguish among types of functions by observing the pattern of change.

Continued

Level 2: Familiarization

At this level students can use an equation to generate ordered pairs in a table and/or plot points on a graph, but this is carried out as a sequence of procedures. However, students can grasp the trend in simple cases, namely, they can recognize when a dependent variable changes with an independent variable (the mutual change), and may be able to characterize it as strictly increasing or strictly decreasing. When a function is not strictly increasing or decreasing, students at this level can visualize how one variable does (or does not) change with the other by describing how the graph changes over different intervals.

At this level students have difficulty with perceiving the connections between the different representations of a function. Because graphing is pointwise, common errors that may be in evidence include (1) seeing only marked points as "part" of a graph; (2) connecting points when a graph is not continuous; (3) neglecting considerations of scale; (4) when interpolating or extrapolating points on a graph, reverting to simple examples, and overgeneralizing linearity; and (5) confusing relationships with pictorial aspects of the situation. Connecting points when a graph is not continuous persists through Level 4.

At this early stage of learning about functions, students are developing a ***scheme*** for the relationship between two sets of numbers, but it is pointwise and based on operations. Students at this level think of a function as an equation or a computational process, and do not perceive a link between alternate representations of the same function, that is, they can work with the equation, the table, or the graph of a function, but not necessarily understand the connections between those different representations. Even equivalent functions (equations) are not necessarily perceived as such when alternate variable names are used, and alternate computational processes (e.g., a function expressed in explicit form or recursive form) are definitely not perceived as the same function. They can evaluate expressions, but do not yet understand formal function notation (e.g., what f and $f(x)$ mean), nor do they conceive of a function as mapping from inputs to outputs.[27] It is important to note that while a formula is one kind of rule for a function, it is certainly not the only kind. At this level, students might still reject a function defined by a graph as such. It is not until Level 3 that students recognize that the rule for a function is not necessarily a formula.

Local Relational—Function as a Formula
Scheme of a relationship between quantities pointwise, bound to a computational process or algebraic expression

Level 1: Preexposure

At this level students can work with sequences by extending them or finding missing terms. They can complete patterns. Although they would not use function notation, they can informally express recursive rules for sequences (e.g., "Keep adding 3"). They would not express rules in explicit form (take the term as input and generate the corresponding value as output). Students at this level can evaluate simple expressions.

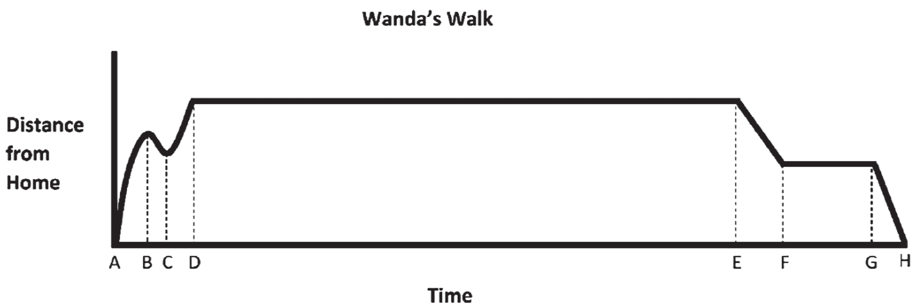
At this level students have difficulty with thinking relationally, namely connecting a pattern in one dimension to a pattern in another.

Students at this level have a solid perception and understanding of patterns in one dimension, attending to quantities and operating on them. At this stage, students had not yet been introduced to functions, nor do they perceive a relation or mapping between two patterns. In the context of proportional reasoning, students may employ scalar strategies but not functional strategies.[28]

Quantity Perception
Perceiving a quantity as is, and operating on it

Appendix B: Complete Tasks

Wanda's Walk



One school day, Wanda walks on a straight, flat path from her home to school and back again. The graph shows Wanda's distance from home with respect to time. For instance if the point (10 am, 1 mile) were on the graph, we could say that at 10 am, Wanda was 1 mile from home.

a. At what point(s) in time is Wanda at home?

____

b. Write a story about Wanda's walk to and from school based on the graph. In particular, explain what Wanda might be doing between each pair of time points below:

A to B: ____

B to C: ____

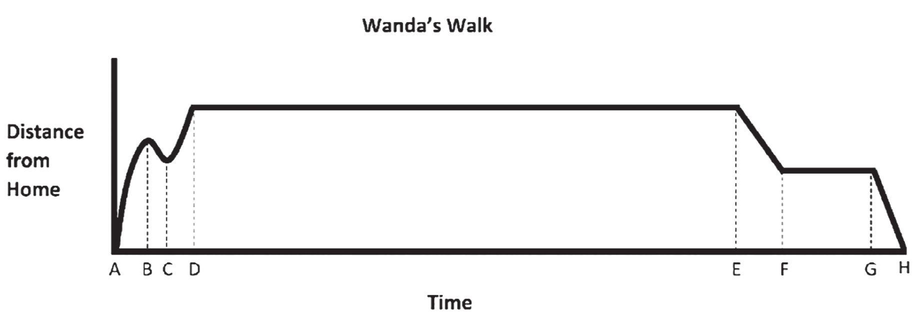
C to D: ____

D to E: ____

E to F: ____

F to G: ____

G to H: ____



c. During which of the following time intervals does Wanda

(1) spend the most time?

☐ A to B ☐ B to C ☐ D to E ☐ G to H

(2) spend the least time?

☐ B to C ☐ D to E ☐ E to F ☐ F to G

(3) have the greatest speed?

☐ A to B ☐ D to E ☐ E to F ☐ F to G

(4) have the least speed?

☐ A to B ☐ B to C ☐ C to D ☐ D to E

d. Does the graph of Wanda's walk to and from school represent a function? Explain in terms of time and her distance from home.

____
____
____
____
____
____

Annika's Bakery

Annika's bakery is closed on Monday. On Tuesday through Thursday, the bakery prepares 3 chocolate fudge cakes each day. On Friday through Sunday, the bakery prepares 5 chocolate fudge cakes each day.

a. Is the number of chocolate fudge cakes prepared each day a function of the day of the week? (Explain).

____
____
____
____
____
____
____
____
____

b. Make a table that shows the number of chocolate fudge cakes prepared on each of the 7 days of the week.

c. Initially, you indicated that the number of chocolate fudge cakes (is, is not) a function of the day of the week. Use the table to support this claim, or to explain why you have changed your mind.

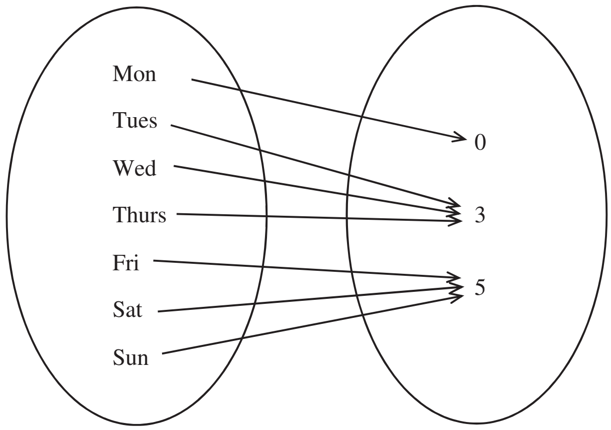
____
____
____
____
____
____
____
____

d. Based on your table, make a graph that shows the number of chocolate fudge cakes prepared on each of the 7 days of the week.

e. You last indicated that the number of chocolate fudge cakes (is, is not) a function of the day of the week. Use the graph to support this claim, or to explain why you have changed your mind.

____
____
____
____

The following arrow diagram shows the number of chocolate fudge cakes prepared on each day of the week at Annika's bakery. On Monday the bakery is closed.



f. You last indicated that the number of chocolate fudge cakes (is, is not) a function of the day of the week. Use the arrow diagram to support this claim, or to explain why you have changed your mind.

Secret Messages

Annika wants to leave a note in the bakery that only her sister Elke (and you!) can read. She decides to *encrypt* her message using a cipher. A cipher is a rule that maps letters to new letters or symbols. The letters in the original message are called *plaintext* and the letters or symbols in the encrypted message are called *ciphertext*. To read the message, the recipient must *decrypt* it by mapping the encrypted letters or symbols back to the original letters. Annika uses the chart below to encrypt her message. Only Annika, Elke (and you!) have access to this chart.

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	B	E	H	K	N	Q	T	W	Z	C	F	I	L	O	R	U	X	A	D	G	J	M	P	S	V	Y

The first row of the chart gives the letters in the plaintext, and the second row gives the corresponding letters of the ciphertext.

Annika encrypts a message by translating the plaintext to ciphertext. Elke decrypts Annika's message by translating the ciphertext letters in the chart back to the plaintext letters.

1) Annika needs your help encrypting her message. She asks you to encrypt the words below by writing the corresponding ciphertext letters next to each word.

SECRET _ _ _ _ _ _
FUDGE _ _ _ _ _

2) Once the note is written, Elke would like your help decrypting it. Help Elke by decrypting the words that have blanks above them.

THE SECRET INGREDIENT _ _ _ _ CHOCOLATE FUDGE CAKE IS _ _ _ _ _ _.
GWN DNHANG ZOTANKZNOG ZO LV HWRHRIBGN QJKTN HBFN ZD HRQQNN.

3) Is the mapping from the plaintext to the ciphertext a function? Is the mapping from the ciphertext to the plaintext a function? Explain.

Annika is concerned the cipher may be broken, so she decides to change it. Here is the new chart she comes up with:

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	E	G	I	K	M	O	Q	S	U	W	Y	A	C	E	G	I	K	M	O	Q	S	U	W	Y	A	C

4) Do you think this chart will work well for encrypting and decrypting messages? Why or why not?

Is the mapping from the plaintext to the ciphertext a function? Is the mapping from the ciphertext to the plaintext a function? Explain.

Suggested citation:

Graf, E. A., Peters, S., Fife, J. H., van Rijn, P. W., Arieli-Attali, M., & Marquez, E. (2019). *A preliminary validity evaluation of a learning progression for the concept of function* (Research Report No. RR-19-21). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12257

Action Editor: Heather Buzick

Reviewers: Isaac Bejar and Heather Howell