Loughborough University

This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

The development of a function concept inventory

PLEASE CITE THE PUBLISHED VERSION

http://dx.doi.org/10.1007/s40753-016-0030-5

PUBLISHER

© Springer

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

O'Shea, A., S. Breen, and Barbara Jaworski. 2019. “The Development of a Function Concept Inventory”. figshare. https://hdl.handle.net/2134/21621.

The Development of a Function Concept Inventory

This paper describes the development of a concept inventory, a test designed to investigate undergraduate students' understanding of the concept of function. A central purpose was to address *conceptual* understanding. We outline a set of elements of the understanding of function, based on key properties of the function concept, which were used to construct test items. We describe the design and validation process for the concept inventory and comment on some implications for the refinement of the instrument and its use.

Keywords: function; concept inventory; conceptual understanding; assessment design.

Introduction

Functions are central to present day mathematics. As Selden and Selden (1992) explain, the function concept has evolved with mathematics and now plays a unifying role. For instance, going beyond calculus, functions are widely used in the comparison of abstract mathematical structures. Yet comprehension of the function concept is remarkably complex and studies have shown that undergraduate students often have difficulties with this concept (Carlson 1998) and even with the notion of variable (Trigueros and Ursini 2003). Pettersson (2012) identified function as being a threshold concept in mathematics; that is, that it is transformative (understanding the concept leads to a new perception of the subject), irreversible (the change in perception is unlikely to be forgotten), integrative (the new understanding reveals connections and relations with other topics), bounded (in the sense that these concepts often lie at the borders of disciplinary areas), and troublesome (in that it presents difficulties to students) as described by Meyer and Land (2003). Since students usually first encounter the concept of function in school, university lecturers might assume that undergraduates have crossed this particular threshold of understanding. However due to the complexity and troublesome nature of the concept, this is often not the case (Pettersson, Stadler and Tambour 2013).

Working with functions in various contexts and in diverse areas of mathematics, requires the ability to think flexibly about functions and to appreciate them not just as actions

and processes but as mathematical objects. The importance of this has been recognised by many authors. Based on Piagetian constructivism, Dubinsky and colleagues (e.g., Briedenbach, Dubinsky, Hawks and Nichols 1992) postulated a hierarchy of concept development, in which the student starts from actions, shifting to processes, developing mathematics objects and ultimately mental schemas. Transitions through the stages of

action -->process-->object-->schema

are usually not linear, but involve shifting between stages through which a concept develops. Dubinsky and McDonald elaborate actions with reference to the concept of function as follows:

> With an action conception of function, for example, an individual may be limited to thinking about formulas involving letters which can be manipulated or replaced by numbers and with which calculations can be done. We think of this notion as preceding a process conception, in which a function is thought of as an input-output machine. What actually happens, however, is that an individual will begin by being restricted to certain specific kinds of formulas, reflect on calculations and start thinking about a process, go back to an action interpretation, perhaps with more sophisticated formulas, further develop a process conception and so on. In other words, the construction of these various conceptions of a particular mathematical idea is more of a dialectic than a linear sequence. (Dubinsky and McDonald 2001 p.277)

They suggest that "An *object* is constructed from a process when the individual becomes aware of the process as a totality and realizes that transformations can act on it" and "a *schema* for a certain mathematical concept is an individual's collection of actions, processes, objects, and other schemas which are linked by some general principles to form a framework in the individual's mind that may be brought to bear upon a problem situation involving that concept" (p.276-277). Sfard (1991) talks of a duality, with processes and objects acting as two sides of the same coin and conceptualisation shifting between them. She suggests a process of reification, consisting of 3 stages of concept construction - interiorization, condensation and reification -- through which processes become objects. In interiorization, operations on lower level mathematics objects enable the learner to get acquainted with processes which will

eventually give rise to a new concept; condensation involves the squeezing of lengthy sequences of operations into more manageable units, thinking of a process as a whole without going into details but attaching a label for a new concept to be born; reification involves an ontological shift, a quantum leap, a sudden ability to see something familiar in a totally new light. A new entity (an object) is soon detached from the process that produced it. Sfard (1991) speaks of reification as a "rather complex phenomenon" (p.30), causing obstacles and frustration for learners, illustrating why reaching an understanding of a function, for instance, as an object can be said to be troublesome though transformative.

Given the importance of function in mathematics, it is useful for a lecturer to have information on their students' level of understanding of this concept. In this paper, we will describe the development of an instrument to elicit this information, namely a function concept inventory. In particular, we were hoping to gain insight into students' understanding of some key properties of a function object.

Literature Review

A *concept inventory* is an instrument, or test, designed to explore *conceptual* or *relational understanding* as opposed to *procedural* or *instrumental competence* (e.g., Skemp 1976; Hiebert and Lefevre 1986). The term *concept inventory* originates in the Physics Education literature. Hestenes, Wells and Swackhamer (1992) developed a concept inventory for the concept of force. Their intention was to explore students' knowledge and understanding about this basic concept. To do this, they decomposed the force concept into six *conceptual dimensions* (similar to what we have described as 'key properties'), and then designed multiple-choice questions to illuminate understanding and misunderstanding in each dimension. Hestenes et al. (1992) claim their inventory is a very good detector of Newtonian thinking and probes commonsense misconceptions. They say their instrument has proven valuable at every level of physics instruction from high school to university, providing sound technical knowledge required for effective teaching. Not only this, but it has been widely used as a pre- and post-test to evaluate gains in student understanding after instruction. Moreover, Epstein (2013) maintains that the

Force Concept Inventory has “spawned a dramatic movement of reform in physics education” (p.1018). Concept inventories have also been used in other subjects such as biology and chemistry (Garvin-Doxas and Klymkowsky 2008; Mulford and Robinson 2002).

Carlson and her colleagues developed the Precalculus Concept Assessment (PCA) to test students’ understanding of function and of rates of change (Carlson 1998; Carlson, Oehrtman and Engelke 2010). In a similar manner to Hestenes et al. (1992), they first developed the PCA taxonomy to articulate foundational knowledge for beginning calculus. Their taxonomy consists of three categories of reasoning abilities and three categories of understandings which they claim are essential for using central concepts of precalculus and understanding key concepts of beginning calculus. The reasoning abilities identified in the PCA taxonomy are: process view of a function (viewing a function as a process instead of an action); covariational reasoning (dealing with change in two related variables); computational abilities. The categories of understandings are: understand meaning of function concepts (such as composition, inverse, rate of change, evaluation); understand growth rates of function types (for example linear, rational, exponential functions); understand function representations (graphical, algebraic, numerical, contextual) (Carlson et al. 2010 p. 120). Carlson and her team then used the taxonomy to design, develop and validate the PCA in a four-phase process. This included reviewing existing research on learning precalculus and beginning calculus, conducting a series of focussed studies on characterizing reasoning abilities and understandings, carrying out clinical interviews to validate questions and distractors for multiple choice items, and using quantitative data from the final version of the PCA to establish the meaning of a PCA score. They explain how the PCA is useful for assessing pre-post learning and thus for comparing various approaches to teaching precalculus courses, and suggest it may have potential as a ‘calculus readiness’ assessment tool.

O’Callaghan (1998) developed a conceptual model to describe the understanding of functions. The elements of his framework stem from applying theory about the sources of meaning in mathematics (following Kaput (1989)) specifically to the function concept

and he described this model in terms of four competencies: modelling (the ability to represent a problem situation using functions); interpreting (the ability to interpret different representations of functions); translating (the ability to move from one representation to another); reifying (the creation of a mental object from a process or procedure) (O'Callaghan 1998 p 25). He then designed an instrument to test these competencies, by attempting to operationalise these abstractions and formulate them in terms of a problem-solving environment, and used it to investigate the effects of different types of instruction on students' knowledge of functions. In fact, the desire to carry out such an investigation motivated his development of a conceptual model.

In our study, we wished to investigate aspects of students' understanding of some key properties of function as an object. Carlson et al. (2010) recognised a 'process view of function' as central to the understanding of precalculus and beginning calculus and included it as one of the three reasoning abilities in their PCA taxonomy. They explain that for many students taking precalculus modules "reasoning is dominated by a static image of arithmetic computation used to evaluate a function at a single numerical value" (p.115). This is problematic as students who are unable to imagine a continuum of input values producing a continuum of output values, that is, conceptualise a function as a process, have difficulty inverting and composing functions, which can in turn hamper their effective use of functions to solve word problems. The PCA attempts to assess students' understandings of the meanings of function concepts such as composition and inverse among others. The four competencies described by O'Callaghan (1998), forming components of his model for function, include 'reifying' which goes beyond the process view of a function. Reification represents the "ultimate goal" or "final stage in the acquisition of function" and is defined as "the creation of a mental object from what was initially perceived as a process" (O'Callaghan 1998 p.25). The research question which accompanies this competency (reification) involves students performing operations (such as composition) on functions and knowing properties of families of functions (such as linear and

quadratic) and thus overlaps with the assessment of the understandings outlined in the PCA taxonomy (understanding meaning of function concepts and growth rate of function type).

Furthermore, it would seem that the questions O'Callaghan poses to assess reifying competence could be successfully completed by students using a process view of function. For instance, one of these questions gives expressions for two functions, $C=0.10(p-1000)$ and $p=100n-n^2$, (O'Callaghan 1998 p.30) and asks the students to find $C(p(50))$ and then an expression for $C(p(n))$, which it would appear could be completed using an operational approach (action or process view) rather than a 'structural' approach (object view). Breidenbach et al. (1992, p.251) contend that students with an action conception of function would be able to perform the steps necessary to find an expression such as that for $C(p(n))$ when given specific formulae for the functions involved. In an attempt to avoid this type of response, we hoped that the inclusion of questions in which a specific formula~~example~~ or description of a function was not given would give better insight into students' understanding of function and this influenced the design of our questions (for instance, Questions 5 and 10 described below).

Another instrument, called the Calculus Concept Inventory (CCI), was developed by Epstein (2013) and his co-workers. The aim of the CCI is to measure conceptual understanding rather than computational skill, with a focus on understanding the concept of derivative, and as such involves material beyond the notion of function itself. The authors report that the first version, drafted in 2005, was too difficult for freshman students and was revised in 2006. The revised test has 22 multiple choice items and has been administered as a pre- and post-test in a variety of universities, predominantly in North America. Epstein (2013) reports that the results suggest that the class performance (measured by a normalized gain) in almost all traditional courses showed little or no improvement from the pre-test to the post-test, however the courses taught using interactive engagement (IE) methods showed significant gains. Epstein defines IE methods as those which involve 'activities which yield immediate feedback through discussion with peers and/or instructors' (Epstein 2013 p. 1020).

While alternative instruments to test understanding of the function concept exist (O'Callaghan (1998), Carlson et al. (2010)) they were not all readily available to us at the time when we designed our concept inventory, nor were they precisely aligned with the elements of understanding we identified as most important for our students. For instance, our own experience of teaching has shown us that students often have difficulty in determining whether a given relationship (in the form of a formula(e), graph, table or verbal description) represents a function or not and so we thought this was an important aspect to have included in a concept inventory. ~~However, neither Carlson et al. (2010) nor O'Callaghan (1998) seem to address this. We also believed that existing instruments would require more time (to administer and correct) than we felt was practical for an instrument which would ideally be available for use as a pre and post-test with each cohort of students. Therefore, we set about designing a new instrument.~~

Development of the Function Concept Inventory

Motivation

The three authors of this paper have each taught courses to first year undergraduate students in which functions play an important role. At the time of this study, we were teaching groups of finance or humanities students attending one Irish university, pre-service primary teachers or humanities students at a second Irish institution, and materials-engineering students in a UK university. In all three cases, we were concerned that our students seemed to come to their university programme with a procedural view of functions, seeing a function merely as a formula relating variables, a machine for finding the output from a given input, or a form of equation, but unappreciative of key conceptual ideas such as uniqueness of image or existence conditions for an inverse function. We wanted to investigate these concerns with a view to gaining insights into how to address them in our programmes. Thus, in an effort to evaluate our students' understanding of the function concept we developed a function concept inventory. We made attempts to seek a pre-existing instrument, but neither that of O'Callaghan (1998) or Carlson et al. (2010) seemed suitable: either they went beyond our focus on functions or they

did not cover all aspects of the concept of function that we wanted to investigate. At the time, we were not aware of the CCI, although retrospectively we see that CCI questions go beyond the material we were addressing. Also, we were concerned that O'Callaghan's instrument, containing 14 constructed response questions with multiple parts would take too long to both administer and correct if it were to be used year after year. Thus we designed our own assessment tool.

Process

The process that we followed in creating the concept inventory is similar to that advocated by the AERA, APA and NCME (1999) in the Standards for Psychological and Educational Testing. In that document, four phases of test development are identified (p. 37):

> Phase 1: Delineation of the purpose of the test and scope of the construct or the extent of the domain to be measured;
> Phase 2: Development and evaluation of the test specifications;
> Phase 3: Development, field testing, evaluation, and selection of the items and scoring guides and procedures; and
> Phase 4: Assembly and evaluation of the test for operational use.

In the next section, we will outline the stages of the development of our inventory.

Theoretical Basis for Elements of Understanding Identified

Following the example of Hestenes et al. (1992), we began by identifying key properties of the function concept with which we were concerned; we did this by drawing on the literature on the subject as well as by drawing on our own experience as teachers. For example, from our own experience we felt that students sometimes struggle with *the difference between a function and an equation*. The literature provided more evidence of this; Vinner (1983) investigated students' concept definitions and concept images of the notion of function and found that one of the major categories of definition was that a function is an algebraic term, a formula, or an equation (p. 300). Sajka (2003) also found that the concept of function is closely related to that of equations in some students' minds (p. 238). If we think of the concept of functions from a perspective of reification, we might see a transition from an incomplete conceptualisation of

functions and equations, in which their important difference is only vaguely perceived, towards a recognition of them as different conceptual objects related in important ways.

Carlson et al. (2010) and O'Callaghan (1998) recognised the value of being able to work with *different representations of functions*, with O'Callaghan noting that two of the most common representational systems for functions to this day are graphs or tables. This is perhaps not surprising considering the historic development of the function concept can be traced from tables to curves and on to the formal definition in analysis (Balacheff and Gaudin 2010). The work of Vinner (1983) and Vinner and Dreyfus (1989) highlighted the difficulties that students have with defining functions and being able to *classify a relationship as being a* function or not. Slavit (1997) argued that a property-oriented view of function can help students appreciate functions as objects. In addition, Carlson et al. (2010) included understanding the meaning of function concepts or *properties* as one of their dimensions of understanding. The ability to *use functions in context* or as part of a mathematical model has been recognised as important in the literature, for example O'Callaghan (1998) cited modelling and interpreting as two of his four competencies in this area. He explained these as formulating a mathematical representation of a problem situation and reversing this process. Covariational reasoning, or the ability to determine how the output values of a function are changing by imagining changes to the input values, was included by Carlson et al. (2010) as one of their core reasoning abilities and this resonated with our experience as lecturers.

Design of a set of elements of understanding

Based on these insights we suggested a set of elements of understanding which we should like students to develop, enabling them to work with key properties of functions. The aim of mathematics teaching at university level should be to give students tools and opportunities to develop their understanding of concepts, that is, where students appear to have only an action and/or process view of functions, to move to the reification or object stage of concept construction. It is difficult to judge whether a person has an object view of function, but based on the literature cited above, we endeavoured to outline some indicators which would

allow us to tell if a student has made some progress on the action/process/object continuum. The six elements of understanding identified are:

1. the ability to distinguish between functions and equations;
2. the ability to recognise and relate different representations of functions and use them interchangeably;
3. the ability to classify relationships as functions or not functions;
4. the ability to have a working familiarity with properties of functions such as one-one/many-one, increasing/decreasing, linearity, composition, inverses;
5. the ability to use functions in context, modelling and interpreting;
6. the ability to engage with co-variational reasoning.

Design of Instrument

Using these elements of understanding, we designed an initial set of fourteen assessment questions. With considerable discussion of what to include, we wrote thirteen of these questions ourselves, drawing on and modifying questions we had used in our courses, and also choosing one item on co-variational reasoning available from the PCA (Carlson et al., 2010). We intended that the test would be administered in class-time and so did not want it to be too long. After further discussion we reduced the number of questions to twelve which spanned all elements of understanding listed above. .

Q1. Let f(x) =3x+5. Given the equation 3a+5=2, which of the following are true? (There may be more than one true statement.)

(a) f(a)=3a+5,	True False
(b) f(a) = 2,	True False
(c) f(x)=2 for all x,	True False
(d) f(x)=2 for some value of x.	True False

Figure 1: Question 1

For example, Question 1 (Figure 1) tests the ability to distinguish between function and equation, seeing these as two distinct but related objects. For parts (b) and (d) students need to make a connection between the function f and the equation 3a+5=2, but in part (c) students need to recognise the difference between the equation and the function. Students may recognise f as a linear function and thus realise it has unique functional values for all real numbers; thereby recognising that f(x) = 2 => x = -1, and that no value of x other than -1 can give a functional value of 2. The question thus tests Element 1 and relates also to Element 4.

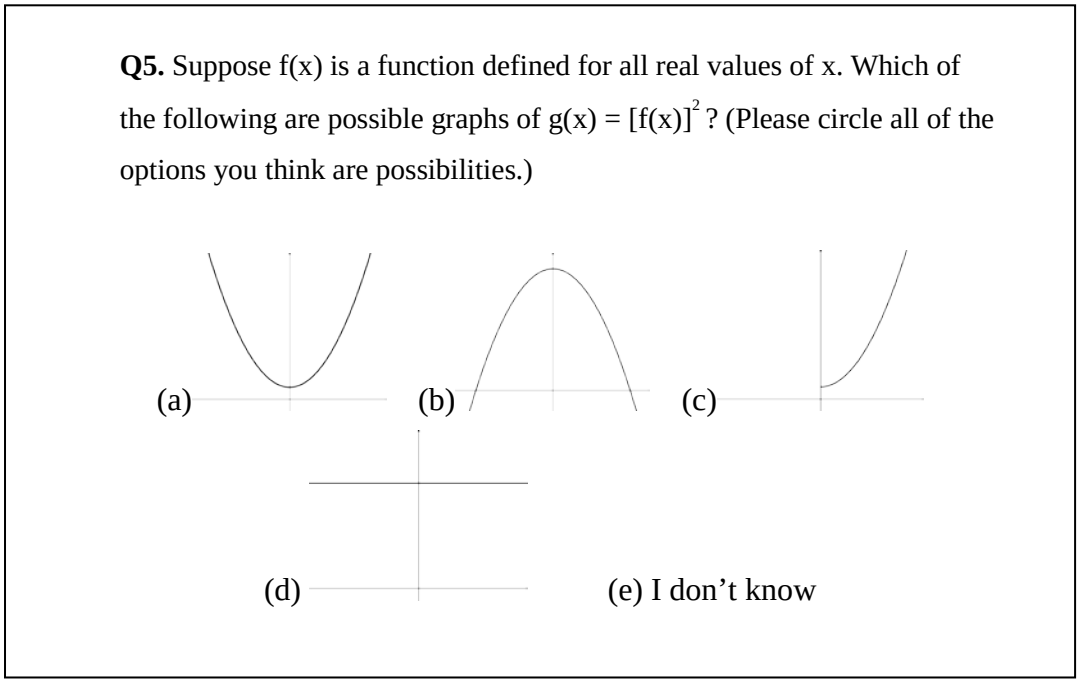


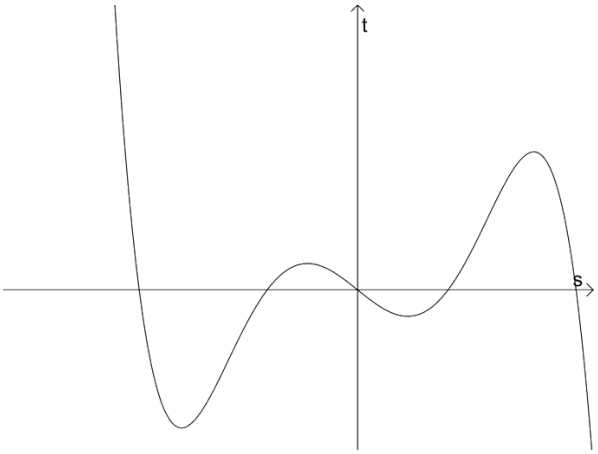
Figure 2: Question 5

Question 5 (Figure 2) tests Element 2 and touches on Element 4. Students need to be able to think in terms of a function defined for all real values and visualise its square. As f has not been specified, students cannot take an operational approach but must focus on structure. They should recognise that (b) does not represent the square of a function, that (c) excludes values for which the square would be defined and that (d) is acceptable for a constant function (something that students frequently ignore). Note that we assume here that students will recognise the horizontal axis in each graph as the line y=0 and therefore that the function graphed in (a) takes no negative values.

Except for Element 6, most of the elements were assessed using more than one question: for example, Element 3 was tested using three questions. These questions are given in Figure 3 (the full concept inventory can be found in Breen, Jaworski and O'Shea (2012)). Questions 8 and 9 were designed to examine the students' ability to decide whether relationships were functions or not when the information was presented in a graphical or tabular form, while Question 10 concerns elements of the formal definition of a function. We see each of these questions as testing the students' ability to work with the key properties of the function concept. In Question 8 students are asked to deduce function relationships from a graphical representation; this requires them to identify relationships which are functions and those which are not. As no formula or analytical expression has been given for the function, it is difficult for the student to take an operational approach in terms of following through a series of actions or a process. Question 9 challenges the view of functions as always defined by formulae: for example, a student with an action or process view of function might look for a formula relating house size to its selling price, whereas a student with an object view might see a function as a relationship between two sets of numbers which is required to have unique outputs. In particular, Question 10 requires students to think in terms of the formal definition of a function allowing (i) but not (ii). Since a specific function is not given here, we hoped to test whether students were able to think about properties of functions without performing actions or calculations or working through a process.

Ten of the twelve questions contained in the inventory were multiple choice. The remaining two were short constructed response questions for which partial credit could be awarded. Such questions are much more difficult to mark, and would probably cause problems when comparing results if the test was to be used widely; however we feel that the high quality of information on students' understanding gleaned from this type of question makes the case for including them in a test compelling. We hope to report on the findings relating to these questions elsewhere. Some of the multiple choice questions had multiple parts, and when sub-questions were counted there were 15 separate items; each was graded separately.

Q8. Consider the curve below which describes a relationship between t and s.



Which of the following is true? (Please circle one of (a)-(d), or (e).)

a) s is a function of t but t is not a function of s.

b) t is a function of s but s is not a function of t.

c) s is a function of t and t is a function of s.

d) s is not a function of t and t is not a function of s.

e) I don't know.

Q9. Let s denote the size of a house and p denote its selling price. The tables below show the sizes and prices of houses sold each month by a suitable estate agent.

May

S	p
$100m^2$	€300,000
$120m^2$	€375,000
$75m^2$	€257,000
$90m^2$	€300,000
$110m^2$	€350,000

June

s	p
$100m^2$	€302,000
$120m^2$	€370,000
$70m^2$	€200,000
$131m^2$	€400,000
$120m^2$	€350,000

July

s	P
$110m^2$	€330,000
$90m^2$	€270,000
$80m^2$	€240,000
$50m^2$	€125,000
$120m^2$	€360,000

For which months is it true that p could be a *function* of s? (Please circle all of the options you think could be functions.)

(a) May (b) June (c) July (d) I don't know

Q10. Suppose f(x) is a function defined for all real values of x. For each of the following statements decide if the statement is always, sometimes or never true.

i. There are two different real numbers a and b such that f(a)=f(b).

Always ☐ Sometimes ☐ Never ☐

ii. There are three different real numbers a, b, c such that f(a)=b and f(a)=c.

Always ☐ Sometimes ☐ Never ☐

Figure 3: Questions 8, 9, and 10.

Validation

Piloting and Refinement

The inventory was piloted in three different ways: it was administered to a group of students attending a bridging course at an Irish university; it was sent to a group of Irish second level mathematics teachers for comment; it was given to university lecturers for comment. Based on the feedback from this piloting process, some of the wording of the questions was altered and one question was changed substantially. This question concerned the property of injectivity, a topic with which the Irish mathematics teachers felt their students would not be acquainted. The question was adapted so that it asked about *increasing* functions rather than *injective* ones; we wished to administer the assessment before any instruction on the function concept at university took place and thus it was necessary to make sure that the material was covered by the school syllabi. The authors considered sample answers to the partial credit items and together developed an agreed marking scheme.

Rasch Analysis

Following the pilots, the concept inventory was administered to three groups of first year students in October 2011: a group of 53 first year engineering students in a UK university studying a basic mathematics course; a group of 37 BA and BEd students in one Irish university, and a group of 127 BA and Finance students in a second Irish university. All of the students in Ireland had chosen to study Mathematics and were taking a first Calculus course.

The test was taken during class time at the beginning of the students' first semester at university before the topic of functions was covered in any of the three modules. Ethical approval for this study was sought and received prior to administration of the concept inventory. The students were told that participation was voluntary and were given 30 minutes to complete the inventory. The scripts were marked by research assistants and the data was compiled into a single file.

When designing a test such as this one, it is important to consider whether it is valid and reliable. Validity here refers to the extent to which the concept inventory measures the variable that it is intended to measure (Gravetter and Furzano 2012, p. 78). In our case, we would want to know if our test instrument measures the trait of conceptual understanding of functions. An important aspect of validity in this is content validity (Sireci, 2007), that is the extent to which subject experts agree that the test concerns the concept in question and covers all aspects of this concept. As noted previously, we piloted the instrument in various ways, with students, secondary teachers and with university mathematicians. The experts in this pilot phase (i.e. the teachers and mathematicians) did not raise any questions as to the content validity of the concept inventory. A further question regarding the validity of the concept inventory is whether the test items combine to give a measure of one single trait i.e., that of conceptual understanding of function. In our design phase we outlined the different aspects or elements of this type of understanding using the literature and our own experience. We designed the test items based on these elements, and then it was necessary to check whether these items actually measured a single construct as set out above: that is, whether the six elements described were contributing to an understanding of the concept of function. Note that this scenario is common in test construction. For example, when measuring mathematical literacy, the PISA studies use items that concern different mathematical content areas as well as thinking processes but together form an instrument which measures the underlying construct (OECD 2014).

In order to further study this aspect of the validity of the concept inventory we used Rasch Analysis (Bond and Fox 2007) by means of the computer software Winsteps (Linacre 2009). The Rasch model is an Item Response Theory model which can be used to evaluate tests, especially those that claim to measure one construct. This is for example, how the PISA studies validate their test instruments and construct measures of mathematical literacy (OECD 2014). Similarly, Wilson and MacGillivray (2007) used Rasch Analysis to validate a test designed to measure the basic mathematical skills of first year university students. The test, in their case, consisted of items relating to different components of algebraic proficiency. Furthermore,

Pantziara and Philippou (2012) used the Rasch model to investigate the validity and reliability of a test to measure understanding of fractions; we will use similar methods here.

The Rasch model is based on the assumption that useful measurement involves the consideration of a single trait or construct at a time (i.e. assumption of unidimensionality), and it incorporates a quality control mechanism using error estimates and fit statistics to verify this. For an introduction to this model, please see Edwards and Alcock (2010). The Rasch analysis computes weighted and un-weighted mean square statistics (called the infit and outfit statistics respectively) for each item. These are chi-square statistics divided by their degrees of freedom and thus have expected values of 1. Bond and Fox (2007 p.243) report that a reasonable range of infit and outfit statistics for test items is 0.7 to 1.3. We used the dichotomous Rasch model; that is we graded each multiple choice item on our test as either correct or incorrect. Fit statistics for all items were computed, with infit statistics ranging from 0.85 to 1.18 and outfit statistics between 0.74 and 1.29. Thus, all items are shown to be behaving well and contributing to the measurement, providing evidence that the items on the test are working together to measure a single construct. (Full details of the item infit and outfit statistics can be found in Table 1.) In addition, the point-measure correlation (equivalent to the point biserial correlation) was computed for each item. This measures the correlation between scores on an item with the average scores on the remainder of the test. Wolfe and Smith (2007 p 206) recommend that these correlations should ideally be above 0.3, but report that for a test such as a concept inventory any positive correlation is acceptable. There seems to be a difference of opinion in the literature about this cut-off with Jarrett et al. (2012) quoting various sources who recommend that point biserial correlations should be above 0.2. In our case, all of these correlations were positive, with only those of Q9 (point-measure correlation =0.2) and Q3b (point-measure correlation =0.23) lying significantly below 0.3 and none lying below 0.2. This provides some further evidence that the items are relatively consistent with each other.

Question Number	Measure	Infit MNSQ	Outfit MNSQ	Point Measure Correlation
1(a)	-1.39	1.07	1.18	0.33
1(b)	-0.71	1.05	1.12	0.34
1(c)	-1.44	0.86	0.75	0.53
1(d)	-0.62	0.97	0.93	0.42
3(a)	-1.33	1.01	1.03	0.39
3(b)	-1.31	1.18	1.29	0.23
4	0.15	1.08	1.17	0.29
5	2.5	0.92	0.74	0.29
6	-.11	0.87	0.84	0.49
8	1.98	0.97	0.96	0.27
9	1.9	1.06	1.05	0.2
10(i)	-0.28	1.03	1.02	0.35
10(ii)	0.32	0.97	0.95	0.38
11	-0.47	0.91	0.85	0.48
12	0.8	1.03	0.99	0.32

Table 1: Fit statistics and measures of multiple choice questions

In order for a test to lead to a useful measure, it is important that it is able to discriminate between students with high and low levels of the trait in question (Wolfe and Smith 2005). For example, if all questions are very easy or all are very hard then most students will have similar scores and the test will not give good information. Rasch Analysis computes item difficulties for each question (using the number of correct answers in the sample) and person measures for each participant (using the number of correct answers for that person). These measures and difficulties are given on the same scale. Therefore to have a good test it is important to have a range of item difficulties and also to have this range correspond to the range

of person measures. If there is a significant difference in the range of person measures and item difficulties the test may be too easy or too hard for the group in question. The results of the Rasch analysis using the responses to the multiple choice questions showed that there was a good spread of item difficulties (see also Table 1 below), and an item-person map showed that the range of item difficulties matched well on the whole with the range of individuals' scores. Questions 5, 8 and 9 have measures that are located above the highest person measure for this group. Since this group consists of students at the beginning of first-year in university, it is desirable that some items on our test are difficult for these students. This allows the concept inventory to be responsive to increases in the understanding of the function concept if it was used as a pre and post-test. Therefore, the assessment instrument seems to be appropriate for this group.

As well as being a valid test, we need to examine whether our concept inventory is reliable or not. In this context, reliability refers to the stability or consistency of the measure (Gravetter and Furzano 2012 p. 85). According to Adams and Wieman (2011 p. 1303), there are three main methods of testing reliability of an instrument such as ours. The first is to create different versions of the test and administer them; they point out that this is very time-consuming. An alternative would be to administer the same test to the same group at a later point in time; this was not possible for us since our courses covered the notion of function in some detail and thus the students' performance on a re-test would be affected. The second method advocated by Adams and Wieman (2011) is to administer the same test to equivalent populations and compute a stability coefficient. In our analysis, we randomly split our sample in two parts and computed the resulting item difficulties for each sample. The correlation between the item measures for the two parts was 0.971. The third is to consider internal consistency measures. We used the Rasch model to do this. To begin with, we looked at the item reliability index. According to Bond and Fox (2007 p.311) this index estimates 'the replicability of item placement within a hierarchy of items along the measured variable if these same items were to be given to another sample of comparable ability'. This index is given on a scale running from 0 to 1. The item reliability here was 0.98, which indicates that the order of item difficulty is

unlikely to change. Similarly, we computed the person reliability index (which is analogous to the Cronbach alpha coefficient and measures how robust the person ordering would be if a similar test was used with the same group of students). In our case the person reliability was 0.45 and the Cronbach alpha coefficient was 0.531. We removed some items from the test (namely Q3b and the Q9) and recomputed the person reliability and Cronbach alpha coefficients but this did not lead to significant increases. The data indicates that the concept inventory would not be a good high-stakes test but it does allow us to discriminate between students with high and low levels of understanding of function. Indeed Adams and Wieman (2011 p.1304) state that low values of Cronbach's alpha on a concept inventory are acceptable given the restrictions on the size of such a test and the possible uses of the resulting data. They also posit that high values of such an index could indicate redundant questions. Jarrett et al. (2012) quote Miller (1995) who said that for instruments such as concept inventories tests of internal consistency, like the person reliability index or the Cronbach alpha coefficients, could under-estimate reliability.

The maximum possible score (for the whole test including the partial credit questions) was 26 marks. A marking scheme is available (Breen et al. 2012). The mean score for the group was 12.9862, with a standard deviation of 4.33171. The median score was 13 marks. No student achieved a perfect score; in fact the highest personal score was 23 marks while the lowest was 0. Table 1 shows the (multiple-choice) questions and parts of questions and their difficulty measures. Note that low measures correspond to easy items and high measures correspond to difficult ones: questions 5, 8, and 9 were the hardest questions for the students tested, while questions 1 and 3 were the easiest questions (note that this remained the case for each subgroup when the group was randomly divided into two halves). When computing these measures and the fit statistics, we graded the responses as right or wrong and did not award partial credit.

Discussion

Our aim in this paper was to describe the design of a function concept inventory. As function is a pivotal concept in mathematics and teaching for understanding is a common goal of lecturers,

it was believed that a concise valid and reliable instrument exploring conceptual understanding of functions would be of considerable value, not only to ourselves but to the wider community.

The design began with the identification of elements of understanding of key properties of the function concept. This was related directly to ideas of action/process/object as set out above. We drew on the literature to recognise that an object perspective was desirable while recognising the partial steps on the way to this as indicated by Dubinsky and colleagues (e.g. Breidenbach et al. 1992; Dubinsky and McDonald 2001) and by Sfard (1991). Thus the questions as a whole, while addressing key properties of functions, as indicated in the literature, were designed to provide insight into students' understanding of function. We then wrote questions related to these elements and pilot-tested them. Our questions sometimes concerned more than one element (for instance, Question 9 addressed Elements 3 and 5); this was deliberate on our part since we wanted the questions to span the types of understanding we had identified as important, we wanted each element to have more than one question associated to it if possible, and we did not want the test to be too long. We endeavoured to present a number of questions in such a way that it would be difficult for students to adopt an action or process approach in response, in an attempt to ascertain whether the student has made some progress on the action/process/object continuum.

Epstein (2013) reports that the results for 250 students who completed a pilot CCI assessment prior to undertaking Calculus I appeared to be mostly at the random guess level. The same could be said of the students' performance (collectively) on some but not all of the questions on our concept inventory instrument as reported in Table 1. However Table 1 shows that the concept inventory does have a good range of item difficulties. The Rasch analysis also supported this finding. Adams and Wieman (2011) comment on the necessity of a test like this to have such a range, since it allows discrimination between students and gives the researchers information about the levels of mastery of the concept in question. In addition, if the questions are not accessible to a student before studying the course, they remark that the test would only be appropriate as a post-test. This was why we trialled our test with pre-university students and

secondary teachers, since we wanted to be sure that all items concerned material familiar to school-leavers.

The majority of students who completed the inventory had problems working with functions in real-life contexts. It would be interesting to run the concept inventory test again in the next few years in Ireland, since the second level curriculum has recently been revised and now has as its aim the teaching of mathematics 'in contexts that allow learners to see connections within mathematics, between mathematics and other subjects and between mathematics and applications to real-life' (NCCA 2012 p.6). The students in this study would not have had the benefit of working with the new curriculum; therefore, it would be interesting to investigate whether students who have studied mathematics with a more applicable focus would do better on the concept inventory. It would also provide a means of attempting to measure a change in the implemented curriculum.

From our experience of administering this concept inventory, we would make some changes to the test. Consider, for example Question 3 given in Figure 4.

Q3. Suppose f(x) is an increasing function.

a) Is it true that 3f(x) must be an increasing function?

Yes ☐ No ☐

b) Is it true that f(x)+3 must be an increasing function?

Yes ☐ No ☐

Figure 4: Question 3

Reflection on our results shows that the correct answer to part (b) of this question may have been chosen for invalid reasons– by this we mean that students may have assumed the '+3' term guaranteed f(x)+3 was increasing. The question may have discriminated better or elicited responses more reflective of the students' understanding of the 'increasing' property if it had considered 'f(x)-3' instead. The latter would have challenged the direct link between adding a

value and the concept of increasing, and required a more object-like consideration of the nature of this function. We would also label the axes in Question 5 in order to avoid confusion.

Before modifying our concept inventory for future use, we would first like to evaluate it using think-aloud interviews, as other researchers (see for example Epstein 2013; Carlson et al 2010; O'Callaghan 1998) have done. This method of evaluation would allow us to delve deeper into the type of understanding required to attempt each of our questions and the reasoning behind the students' choices of responses.

Many universities already administer diagnostic tests to students in first-year mathematics modules and then provide support for those students who do not do well on these tests. The diagnostic tests usually focus on basic mathematical skills; however the mathematical community places high value on conceptual understanding and concept inventories could also be used as a diagnostic tool. Hestenes et al. (1992) suggest their Force Concept Inventory can be used thus to identify and classify misconceptions, with errors made by students being more informative than correct answers to the questions posed. Thus, they claim their inventory is particularly useful for teachers in raising their awareness of the misconceptions among their own students. As teachers, we gained important information from the responses of our students on our function concept inventory; the results showed us that our students' understanding of the function concept was limited. This information has been used when redesigning our courses and in particular when designing learning activities. We are more aware that, since the concept of function is so fundamental to a Calculus course, it can be revisited many times in a module and the students can be given many different opportunities to develop their understanding of the concept (Meyer and Land 2003). Thus, the information from the concept inventory can give a teacher data on the areas where their students had most difficulty, and the teacher can then tailor the module appropriately.

The Function Concept Inventory described here could also be used by researchers or researcher-teachers to evaluate the effects of different types of instruction in promoting

conceptual understanding and, in particular, the elements of understanding underlying the instrument's design. O'Callaghan's (1998) conceptual model for describing an understanding of functions was initially developed as part of a project to test whether students following a 'Computer-Intensive Algebra (CIA)' curriculum would develop a richer understanding of functions that their counterparts following a 'Teaching Algebra (TA)' or more traditional curriculum. The test O'Callaghan developed to probe each of the competencies he had identified (modelling, interpreting, translating, reifying) was used to investigate whether CIA students were more competent than TA students when working with functions in each of these ways. Hestenes et al. (1992) also used the Force Concept Inventory to test the effectiveness of pedagogical techniques in both school and university settings. They claim they have abundant evidence that their inventory is a very accurate and reliable instrument for evaluating instruction. However, Caballero et al. (2012) point out that challenges arise when concept inventories are used to make comparative evaluations of curricular course reforms if core course content is affected by the reforms.

In particular, we were intrigued by Epstein (2013) reporting on results obtained on the CCI following the use of interactive engagement (IE) methods between pre- and post-applications of their inventory. Perhaps the teaching of O'Callaghan's CIA curriculum would have been considered to include IE methods. It would be interesting to see if the use of IE methods similarly improved students responses to the Function Concept Inventory presented here. This is certainly an area which begs further research.

Furthermore, the Function Concept Inventory may also be useful as a placement assessment in institutions in which there is a choice of precalculus and first calculus/analysis courses available to undergraduate students. Diagnostic tests have been used to date in Irish institutions in the placement of students (Burke, Mac an Bhaird and O'Shea 2012) but, depending on the programme of study followed by the students concerned there are cases where a concept inventory would be more appropriate.

Acknowledgements

The authors would like to acknowledge the support of a NAIRTL (National Academy for Integration of Research, Teaching and Learning) grant for this research.

References

Adams, W.K. and Wieman, C.E. (2011). Development and validation of instruments to measure learning of expert-like thinking. International Journal of Science Education, 33, 1289-1312.

AERA (American Educational Research Association), APA (American Psychological Association) and NCME (National Council on Measurement and Education) (1999). *Standards for educational and psychological testing*. Washington DC.

Balacheff, N. and Gaudin, N. (2010). Modelling students' conceptions: The case of function. *Research in Collegiate Mathematics Education VII, Conference Board of the Mathematical Sciences, Issues in Mathematics Education*, 16, 207-234.

Bond, T. G. and C.M. Fox. (2007). *Applying the Rasch model – Fundamental measurement in the human sciences* (Second edition). New Jersey, Lawrence Erlbaum Associates.

Burke, G., Mac an Bhaird, C., O'Shea, A. (2012) 'The impact of a monitoring scheme on engagement in an online course.'. *Teaching Mathematics and Its Applications*, 31:191 - 198.

Breen, S., Jaworski, B., and O'Shea, A. (2012). Concept Inventory for Functions. Available from: http://staff.spd.dcu.ie/breens/documents/ConceptInventoryforFunction.pdf [Accessed 19 February 2015].

Breidenbach, D., Dubinsky, E., Hawks, J. and Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics*, 23 (3), 247-285.

Caballero, M.D., Greco, E.F., Murray, E.R., Schatz, M.F., Bujak, K.R., Marr, M.J., Catrambone, R. and Kohlmyer, M.A. (2012) Comparing large lecture mechanics curricula using the Force Concept Inventort: A five thousand student study. *American Journal of Physics, 80(638).*

Carlson, M. (1998). A cross-sectional investigation of the development of the function concept. *Research in Collegiate Mathematics Education III, Conference Board of the Mathematical Sciences, Issues in Mathematics Education*, 7(2), 114-162.

Carlson, M., Oehrtman, M. and Engelke, N. (2010). The precalculus concept assessment: A tool for assessing students' reasoning abilities and understandings. *Cognition and Instruction*, 28(2), 113-145.

Ding, L. & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics - Physics Education Research*, 5, (020103), 1-17.

Dubinsky, E. and McDonald, M.A., 2001. APOS: A Constructivist Theory of Learning in Undergraduate Mathematics Education Research In: D. Holton ed., The Teaching and Learning of Mathematics at University Level. Dordrecht: Kluwer Academic Publishers, pp.275-282.

Edwards, A. & Alcock, L. (2010). Using Rasch analysis to identify uncharacteristic responses to undergraduate assessments. *Teaching Mathematics and Its Applications*, 29, 165-175.

Epstein, J. (2013). The calculus concept inventory – Measurement of the effect of teaching methodology in mathematics. *Notices of the American Mathematical Society*, 60 (8), 1018-1026.

Garvin-Doxas, K and Klymkowsky, M.W. (2008). Understanding randomness and its impact on student learning: Lessons learned from building the biology concept inventory (BCI). *CBE Life Sciences Education*, 7 (2), 227-233.

Gravetter, Frederick J.; Forzano, Lori-Ann B. (2012). *Research Methods for the Behavioral Sciences* (4th ed.). Belmont, CA: Wadsworth.

Hestenes, D., Wells, M. and Swackhamer, G. (1992). Force concept inventory, *The Physics Teacher*, 30, 141-158.

Hiebert, J., and Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jarrett, L., Ferry, B., Takacs, G. (2012). Development and validation of a concept inventory for introductory-level climate change science. *International journal of Innovation in Science and Mathematics Education*, 20(2), 25-41.

Kaput, J. J. (1989) Linking representations in the symbol systems of algebra. In S. Wagner & C. Kieran (Eds.) Research issues in the learning and teaching of algebra (pp.167-194) Reston, VA: National Council of Teachers of Mathematics; Hillsdale, NJ; Lawrence Erlbaum Associates.

Linacre, J. M. (2009) Winsteps (Version 3.68) [Computer Software]. Chicago: Winsteps.com.

Mulford, D. .A. and Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education*, 79 (6), 739-744.

Meyer, J.H.F. and Land, R. (2003). *Threshold concepts and troublesome knowledge: linkages to ways of thinking and practicing*. Edinburgh: ETL Project (Occasional Report 4).

Available from: http://www.etl.tla.ed.ac.uk/docs/ETLreport4.pdf [Accessed 20 June 2013].

Miller, M.B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modelling. *Structural Equation Modelling :A Multidisciplinary Journal*, 2(3), 255-273.

NCCA, (2012). Leaving Certificate Mathematics Syllabus. Available from http://www.ncca.ie/en/Curriculum_and_Assessment/Post-Primary_Education/Project_Maths/Syllabuses_and_Assessment/LC_Maths_for_examination_in_2014.pdf [Accessed 8 June 2014].

O'Callaghan, B.R. (1998). Computer-intensive algebra and students' conceptual knowledge of functions. *Journal for Research in Mathematics Education*, 29(1), 21-40.

OECD, (2014). PISA 2012 Technical Report. Paris.

Pantziara, M. and Philippou, G. (2012). Levels of students' 'conception' of fractions. *Educational Studies in Mathematics*, 79, 61-83.

Pettersson, K., (2012). *The Threshold Concept of a Function - A Case Study of a Student's Development of Her Understanding*. Sweden: MADIF-8. Available from: http://www.mai.liu.se/SMDF/madif8/Pettersson.pdf [Accessed 26 April 2013].

Pettersson, K., Stadler, E. and Tambour, T. (2013) Development of students' understanding of the threshold concept of function. Paper presented at CERME8. Available at www.cerme.org [Accessed 26 April 2013].

Sajka, M. (2003). A secondary school student's understanding of the concept of function – A case study. *Educational Studies in Mathematics*, 53, 229-254.

Selden, A. and Selden, J. (1992) Research perspectives on conceptions of functions: Summary and overview. In G. Harel and E. Dubinsky (Ed.s) The Concept of Function: Aspects of Epistemology and Pedagogy (pp.1-16): MAA Notes Vol. 25. Mathematical Association of America.

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.

Sireci, S. (2007). Content Validity. In N. J. Salkind and K. Rasmussen (Eds). Encyclopedia of Measurement and Statistics. (pp. 182-184). Thousand Oaks, CA: Sage Pubications. Relational

Skemp,R. (1976).Understanding and Instrumental Understanding, *Mathematics Teaching*, 77, 20-26, Dec 76

Slavit, D. (1997). An alternate route to the reification of function. *Educational Studies in Mathematics*, 33, 259-281.

Trigueros, M. and Ursini, S. (2003). *First year undergraduates' difficulties in working with different uses of variable*, CBMS Issues in Mathematics Education, 12, 1-29.

Vinner, S. (1983). Concept definition, concept image, and the notion of function. *International Journal of Mathematical Education in Science and Technology*, 14(3), 293-305.

Vinner, S. and Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for Research in Mathematics Education*, 20(4) , 356-366.

Wilson, T.M. and Macgillivray, H.L. (2007). Counting on the basics: mathematical skills among tertiary entrants. *International Journal of Mathematical Education in Science and Technology*, 38(1), 19-41.

Wolfe, E.W. and Smith, E.V. (2005). Instrument development tools and activities for measure validation using Rasch models: Part II – Validation Activities, *Journal of Applied Measurement*, 8(2), 204-234.