

Secondary Students' Implicit Conceptual Knowledge of Algebraic Symbolism. An Exploratory Study through Problem Posing

Elena Fernández-Millán^a, Marta Molina^a

^aUniversidad de Granada, SPAIN

ABSTRACT

Through the task of problem posing, we inquire into students' conceptual knowledge of algebraic symbolism developed in compulsory secondary education. We focus on identifying the characteristics of equations and systems of equations that made the problem posing task difficult for the students and analyzing the meanings that they gave to the operations contained in the expressions. To collect the data we used two questionnaires in which students were asked to pose problems that could be solved by using given equations or system of equations. In the second questionnaire a specific meaning for the unknowns in the given expression was suggested. The results complement those of a previous study. Students evidence a good conceptual knowledge of algebraic symbolism when meanings for the unknowns are suggested. Decimal numbers and an equation including parenthesis and multiplication of unknowns are the main elements that made some weaknesses in students' knowledge to surface. The results are more promising. They suggest the potential for compulsory algebra instruction to develop students' conceptual knowledge, although greater attention should be paid to the semantic aspects of algebra if students are to access such knowledge unaided.

KEYWORDS

Algebraic symbolism; conceptual knowledge; equations; Problem posing; unknown

ARTICLE HISTORY

Received 10 August 2017
Revised 28 September 2017
Accepted 15 November 2017

Introduction

For years research on students' understanding of algebra has focused on their procedural knowledge, normally defined as the command of a sequence of steps or actions that may help solve problems (Crooks and Alibali, 2014; Ross and Willson, 2012). Recent decades have brought a change in the approach to researching algebra teaching instruction, however, geared to conceptual understanding as well to determine in greater depth not only the steps followed

CORRESPONDENCE Marta Molina ✉ martamg@ugr.es

© 2017 E. Fernández-Millán & M. Molina

Open Access terms of the Creative Commons Attribution 4.0 International License apply. The license permits unrestricted use, distribution, and reproduction in any medium, on the condition that users give exact credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if they made any changes. (<http://creativecommons.org/licenses/by/4.0/>)

by students to solve problems, but also their understanding of the concepts implicit in the solution. Attendant upon this new approach has been a change in mathematics instruction in which curricular documents explicitly address the need for students to master both procedural and conceptual algebraic knowledge (Crooks and Alibali, 2014; Ross and Willson, 2012).

This new outlook stems from the realisation of the importance of conceptual knowledge (Crooks and Alibali, 2014; Rittle-Johnson and Schneider, 2015; Ross and Willson, 2012) and the shortcomings repeatedly detected in that regard in studies exploring students' algebraic competence, especially around the use of algebraic symbols (Fillooy and Rojano, 1989; Küchemann, 1981; Furinghetti and Paola, 1994; Booth, 1984; Filloy, Rojano and Puig, 2008). The persistence of errors throughout several years of algebra instruction is striking (Álvarez and Gómez-Chacón, 2015; Molina, Rodríguez-Domingo, Cañadas and Castro, 2016).

These two circumstances informed the present study on the conceptual understanding of algebraic symbolism acquired by students as a result of mathematics instruction delivered throughout compulsory secondary education (abbreviated 'ESO' in Spanish). To narrow the field to be covered, the research focused on linear and quadratic equations with one unknown and systems of linear and quadratic equations with two unknowns. The meaning of literal symbols was restricted to the unknown only, for it is what last year ESO students are most familiar with, inasmuch as the instruction received fixes on such meanings.

The task assigned, to pose problems that could be solved by using certain symbolic expressions, was chosen on the grounds of prior evidence (e.g. Lin, 2004, Mestre, 2002; Sheikhzade, 2008) of the utility of problem posing for assessing students' mathematical skills. It was used in an earlier exploration of students' conceptual knowledge of algebraic symbolism (Fernández-Millán and Molina, 2016) that identified the characteristics of the algebraic equations deployed (linear and quadratic equations with one unknown and systems of two equations with two unknowns in which the coefficients, independent terms and solutions were integers) that rendered problem-posing difficult. The meanings attributed to the unknown by students and the operations contained in such equations were also studied. This second study sought to confirm the findings of the earlier research with a new sample of students. It also aimed to delve further into last-year ESO students' conceptual understanding of algebraic symbolism. For that reason, two forms of problem-posing were postulated, free and semi-structured, and the task variables characterising the equations and systems of equations considered were broadened. According to Stoyanova and Ellerton (1996), a problem-posing situation is considered as free when students are asked to generate a problem from a given contrived or naturalistic situation. It is referred to as semi-structured when students are given an open situation and are invited to explore the structure of that situation and to complete it by applying knowledge, skills, concepts, and relationships from their previous mathematical experiences.

Conceptual Understanding of Algebraic Symbolism

Many studies have addressed the distinction between conceptual and procedural knowledge in mathematics. The objectives pursued include reaching

a consensus on their definition, determining which should prevail in mathematics instruction and how they can best be evaluated (Castro, Prat and Gorgorió, 2016).

The pervasive use of the terms conceptual and procedural can be attributed to Hiebert and Lefevre (1986). These authors' characterisation of the two types of knowledge is applied throughout this article. Conceptual knowledge is based on a dense network of relationships among pieces of information that allow flexibility in information access and use. Procedural knowledge consists in two parts: the first is the system of symbolic representation in mathematics and the second the algorithms or rules used to perform mathematical tasks. More recent studies have confirmed that the aforementioned definition of conceptual knowledge as densely interconnected knowledge and the description of procedural knowledge as the ability to perform actions in sequence to solve problems continues to be valid (Rittle-Johnson and Schneider, 2015).

Despite the dilemma posed in research papers over whether procedural or conceptual knowledge should prevail in mathematics instruction, the consensus opinion is that the two should go hand-in-hand. The benefits attributed to conceptual knowledge include support for decision-making about the most suitable procedure for a given situation, a more flexible approach to problem-solving and evaluation of the solution (Crooks and Alibali, 2014).

Conceptual knowledge can be evaluated using indicators for explicit or implicit conceptual knowledge (Castro, Prat and Gorgorió, 2016; Crooks and Alibali, 2014). For the former, the aforementioned authors cited concept definition. Implicit conceptual knowledge evaluation can be broached through the evaluation, judgement, justification and application of procedures (Castro, Prat and Gorgorió, 2016). For instance, in the specific case of the mathematical concepts equivalence (idea that the two members of an equation represent the same quantity), inversion (idea that inverse operations of the same quantity in an equation do not alter the initial value) and cardinality (ability to count), Crooks and Alibali (2014) suggested a variety of tasks to evaluate implicit conceptual knowledge. These include determining whether an operation is correct, reproducing the structure of an equation or operation viewed previously, identifying equivalent equations and interpreting the solution or explaining the procedure used to solve problems. The authors also stressed the importance of specifically identifying and measuring conceptual knowledge and the need for validated tools for its evaluation.

Rittle-Johnson and Schneider (2015) and Ross and Willson (2012) defined the translation between different representation systems as a method for analysing the implicit conceptual knowledge acquired by students¹. That idea has been endorsed by studies in which different representation systems are used to favour the development of conceptual knowledge in algebra (Cedillo, 2001; Charpell, 2001; Ferruchi, Kaur, Carter and Yeap, 2008; Fujii and Stephens, 2008; Ng and Lee, 2009). Translation is the procedure whereby a mathematical object represented by one system of representation is represented in another (Gómez, 2007). This is a cognitively complex process. In addition to

¹ 'Representation system' is understood here to be a structured set of notations, symbols and graphs that, subject to rules and conventions, can be used to express the features and properties of a concept (Castro and Castro, 1997).

understanding the representation systems involved, it calls for the ability to identify and translate to the other system the essential information that defines the concept represented, ignoring superfluous particulars imposed by the representation system in which the concept is expressed (Molina, 2014).

By algebraic symbolism is meant the representation system characterised by the use of written numerals, letters and signs typical of arithmetic and algebra. Algebraic symbolism is a compact and very precise representation system applicable to mathematics as well as other areas. With it, algebraic ideas can be represented independently of the initial specific context in which they arise (Arcavi, 1994) and expressions can be transformed with learned algebraic techniques, irrespective (temporarily) of the meaning of the constituent symbols. Consequently, an essential part of being algebraically competent is the ability to flexibly and opportunistically alternate, on the one hand, the use of actions devoid of meaning and on the other, the pursuit of meanings geared to questioning and choosing strategies, thinking reflectively, connecting ideas, drawing conclusions or formulating new meanings (Arcavi, 2005).

Given the descriptions of conceptual and procedural knowledge adopted in this study, the first dimension of algebraic symbolism referred to by Arcavi is identified with the use of procedural knowledge, and the second with conceptual knowledge. This article focuses on the second dimension, asking students to pose problems that can be solved with given symbolic equations. This task requires translating symbolic to verbal representation and, therefore, it implies identifying quantities and possible relationships among them that may be represented by the starting equations. Letters and operations acquire meaning in a specific context (Wagner, 1981). The aim is to evaluate the implicit conceptual knowledge of the algebraic symbolism involved in each of the equations included in the study.

Review of the Literature

Studies addressing conceptual knowledge of algebraic symbolism do so from different perspectives. A short number of papers discusses teaching strategies or methods that may favour the acquisition of conceptual knowledge of algebraic symbolism. In one, Rittle-Johnson and Star (2007) taught their secondary school students to solve linear equations in three ways: comparing equivalent equations solved using the same method; comparing different types of equations solved with the same method; and comparing different methods for solving the same equation. They found that conceptual knowledge was acquired more effectively by comparing methods than by comparing different types of problems. In a study with secondary school teachers and students, Ross and Willson (2011) analysed the effect of three teaching models on the acquisition of conceptual and procedural knowledge of algebraic symbolism. After analysing the classes delivered by seven teachers, they concluded that the use of symbolic representations and participatory classroom instruction in which different meanings of mathematical ideas were shared by teacher and students helped the latter make connections between their ideas about a given concept, thereby favouring the acquisition of conceptual knowledge. These authors drew attention to the need for more studies on conceptual knowledge in algebra. Chalouh and Herscovics (1988), in turn, geared their research to helping students build conceptual knowledge of algebraic symbolism with models based on the area of

rectangles, while Herscovics and Kieran (1988) deployed arithmetic identities to the same purpose. Both studies found that students interpreted equations correctly more readily than open expressions (with no equal sign).

Along the same lines as addressed in this study, other studies have assessed implicit conceptual knowledge of algebraic symbolism acquired by secondary students. Filloy and Rojano (1989) identified conceptual obstacles in the move from operating with equations with one unknown on one side of the equal sign to operating with equations with unknowns on both sides. To work with this second type of equations students must understand that the expressions in both members are of the same nature and should attribute meaning to the equality of the expressions. That, the authors contended, would call for specific instruction in the context of traditional instruction in algebraic symbolism. Caprano and Joffrion (2006) conducted a study to explore secondary school students' conceptual knowledge of algebraic symbolism with two multiple-choice tasks, in which they were asked to translate verbal to symbolic representation. The conclusion drawn was that the ability to apply existing knowledge to a new situation constituted proof of the acquisition of conceptual knowledge. The authors highlighted the importance of developing conceptual knowledge in mathematics and the relevance of vocabulary in that development.

Some studies that evaluate the conceptual understanding of algebraic symbolism acquired by students focus on literal symbols as components of such symbolism, more than on equations in general. In research on 13- to 15-year-old students, Küchemann (1981) observed that most found it difficult to interpret letters in algebra as unknowns or generalized numbers. Furinghetti and Paola (1994), studying higher education students, found that only a small minority could adequately describe the differences between parameters, unknowns and variables, and most tended to interpret letters as substitutes for objects or words. Both studies concurred with Booth (1984) in identifying the non-equivalent use of letters in arithmetic and algebra as one of the reasons for such difficulties. Filloy, Rojano and Puig (2008) reported cases in which students assigned different meanings to the same letter (for instance, as an unknown and as a variable), when interpreting a single variable equation such as $x + x/4 = 6 + x/4$. In this same vein, Arnau and Puig (2013) reflected on the different meanings (variable vs unknown) that a letter may adopt depending on the semantic field from which the solution to a given problem is broached: functions or equations. The meanings of letters are associated with different algebraic conceptions (Usiskin, 1988). Bills (2001) and Álvarez and Gómez-Chacón (2001), among others, reported that students encountered difficulty in interpreting and distinguishing between the meanings that could be adopted by a letter in a problem and move flexibly from one to another.

Two other groups of authors can be identified who specifically assessed secondary school students' conceptual knowledge using the translation from symbolic to verbal representation as a tool. Resnick, Cauzinille-Marmeche and Mathieu (1987) asked 11- to 14-year-old students to translate algebraic expressions containing additive structures to texts with or without context. They found that none of the students was able to put forward a non-contextualised interpretation (such as a number plus double that number less five) without having previously built a specific story for the given expression. Taking their

data as a basis, they hypothesised that an effective predecessor to learning formal algebra would be to interpret equations with some specific reference.

Molina et al. (2016) analysed the errors made by Spanish 2nd and 4th year ESO students in the non-contextualised translation of algebraic symbolism to verbal language and vice-versa. The errors were classified by the mathematics content involved, distinguishing three categories, associated with: complete/incomplete wording, arithmetic and the characteristics of algebraic symbolism. The last group was sub-divided into: errors in generalising the elements of expressions (translating -4 , for instance, as ‘subtract an even number’), particularisation, assignment of different meanings to the same letter and structural errors. They found that translating symbolic to verbal representation was more accessible than the reverse for both the younger and older students. The most frequent errors in translating algebraic symbolism to verbal language were associated with the characteristics of algebraic symbolism, especially to the last two sub-types. Unlike the other types of errors, whose number declined in the older students, the number of errors related to algebraic symbolism committed by 2nd and 4th year students did not vary significantly. In light of the persistence of some errors with ongoing algebra instruction, the authors suggested the need for more research focusing on the characteristics of algebraic symbolism to acquire a deeper understanding of how students acquire that knowledge.

One generally accepted requirement for successful translations between verbal and symbolic representation is an understanding of unknowns and the mutual dependence described in the verbal wording of the problem, as well as the syntactic characteristics of algebraic symbolism (Kaput, 1989). Those observations inform the distinction between two dimensions in the present analysis of students’ translations: a) the syntactic characteristics of equations and systems preserved by students; and b) the meanings assigned to unknowns and the operations relating such unknowns.

Prior classifications of additive (Carpenter and Moser, 1982) and multiplicative (Castro, 2001) arithmetic problems were used to distinguish between the meanings of additive (addition and subtraction) and multiplicative (multiplication and division) structures. The additive situations defined were change, comparison, combination and equalisation and the multiplicative situations, simple proportionality, comparison and Cartesian product.

An earlier study by the authors (Fernández-Millán and Molina, 2016) identified the characteristics of equations and systems of equations that hindered problem-posing and the meanings attributed by students to the operations and unknowns contained in the equations used. Such characteristics included the presence of more than one unknown, the same unknown on both sides of the equal sign, coefficients greater than two and the multiplication of two or more unknowns. Meanings were more readily attributed to operations involving additive structures, although comparison and equalisation structures appeared in very few of the word problems posed by students, while the Cartesian product was absent in multiplicative structures. The presence of multiplicative structures also heightened the difficulty to attribute meaning to unknowns. Those findings prompted further exploration of the characteristics of

equations that hinder problem-posing and the contexts in which students are liable to use additive and multiplicative structures absent in the earlier study.

Empirical Study

As noted earlier, the research problem addressed in this problem-posing-based study was to analyse the implicit conceptual knowledge of linear and quadratic equations and systems of equations acquired by Spanish ESO students. More specifically, the problem was confined to specific objectives and certain symbolic expressions. The objectives were:

1) to identify and compare the characteristics of equations and systems of equations that hinder students' ability to pose problems, establishing one free and one semi-structured (where a meaning for the unknowns was proposed) situation

2) to distinguish and compare the meanings attributed by students to the operations contained in the equations and systems in the free and semi-structured situations.

The symbolic expressions used were linear and quadratic equations with one unknown and systems of linear and quadratic equations with two unknowns, in which the coefficients, independent terms and solutions were rational numbers. Letters were used to symbolise unknowns.

Participants

The sample, intentionally selected on the grounds of student availability, comprised 32 last year Spanish ESO students enrolled at two schools. Socioeconomic and cultural levels were average in the area where one of the schools was located and low in the other. Both groups of students attended class in a regular basis. Student performance in mathematics was average and heterogeneous in both groups. The results for the two groups were pooled to create a more extensive dataset from a sample with a broader socio-economic spectrum.

The two groups' prior knowledge was theoretically the same. They had been solving equations and related problems from first year, beginning with first-degree equations with one unknown and progressing on to second-degree equations and systems of linear and non-linear equations with two unknowns. When the data were collected for this study they had concluded classroom work on the algebra-related content specified in compulsory education in Spain. More specifically, they had worked with first- and second-degree equations with brackets and denominators and systems of linear and non-linear equations (both with two equations and two unknowns), from both the theoretical and problem-solving perspectives. They had no prior experience in problem-posing.

Questionnaire design

Two questionnaires, labelled 1 and 2, were used to collect the data for this study. Each consisted in seven tasks in which students were asked to pose a problem that could be solved by using the equation or system of equations specified in the task. The equations and systems of equations in the two questionnaires were the same and listed in the same order. The difference

between the two was that in the second, students were furnished a specific meaning for the unknown or unknowns in the equation.

The instructions given to the students for each questionnaire were:

Questionnaire 1: *Write the statement of a problem posed by you that can be solved using the given equation or system of equations and that refers to a context of everyday life.*

Questionnaire 2: *Write the statement of a problem posed by you that can be solved using the given equation or system of equations, taking into account the meaning of the unknowns that is indicated in each case, and that refers to a context of everyday life.*

The symbolic expressions used in this study were designed bearing in mind the three essential factors described below. The first two criteria had been addressed in the earlier study to select the equations set out in the tasks.

1) The structure was to be familiar to students. That involved analysing the units on algebra in secondary school mathematics textbooks (including the book used by students in the year when the data were collected) and identifying the types of equations that prevailed. The equations included on the questionnaires consequently formed part of the ESO mathematics curriculum and had been the object of instruction.

2) Problem-posing was to be feasible. To that end, the equations and systems of equations selected were taken from problem-solving exercises previously performed by students.

3) The findings of the earlier study (Fernández-Millán and Molina, 2016) were taken into consideration. Table 1 lists the equations used in that study and for each, the number of problems posed by students, the number that were correct, the main characteristics of the equations that hindered student problem-posing and the decision to retain them or otherwise in the present study.

Table 1: Equations used in Fernández-Millán and Molina (2016)

#	Equation	No. of problems posed (n=20)	No. correctly posed	Characteristics of equations that hindered problem-posing	Inclusion in present study
1	$8 = x + 6$	19	15 (79 %)	None of significance	No
2	$2x - 1 = 9$	16	8 (50 %)	Coefficient $\neq 1$	No
3	$x + 10 = 6x$	17	6 (35 %)	Unknown on both sides of equal sign Coefficient $\neq 1$	Yes
4	$16 = x^2$	17	8 (47 %)	Multiplication involving two unknowns	Yes
5	$\left. \begin{array}{l} 5x + 3y = 69 \\ x + y = 15 \end{array} \right\}$	15	8 (53 %)	System of equations Coefficients $\neq 1$	No

6	$\left. \begin{array}{l} x + y = 7 \\ xy = 10 \end{array} \right\}$	12	6 (50 %)	System of equations Multiplication involving two unknowns	Yes
7	$20 = x(x + 1)$	13	6 (46 %)	Multiplication involving two unknowns	Yes

The present study introduced a variable that distinguished between integers and decimals as coefficients of unknowns and independent terms. For that reason equations 1 and 2 from the earlier study were eliminated and new equations were introduced with decimals as coefficients and independent terms. System of equations 5 was modified to introduce coefficients different to one in both equations as in the earlier study it was identified as one of the characteristics that hindered problem-posing.

Ultimately questionnaires 1 and 2 contained the same four equations and three systems of equations with single solutions, listed in Table 2 along with the variables studied. The order of the equations relative to the earlier study was varied to determine its possible role in the small number of problems posed in the equations listed in the latter positions.

Table 2: Characterisation of equations and systems of equations used in the study

#	Equation	No. of unknowns	No. of members with unknowns	Coefficient of unknown and independent term	Operation with unknown
1	$10.5x + 2 = 12.5$	1	1	Decimal	Addition with known quantity Multiplication with known quantity
2	$x + 10 = 6x$	1	2	Integer	Addition with known quantity Multiplication with known quantity
3	$x(x + 1) = 20$	1	1	Integer	Addition with known quantity Multiplication with unknown quantity
4	$\left. \begin{array}{l} 5x + 2y = 290 \\ 2x + 5y = 200 \end{array} \right\}$	2	1	Integer	Addition with unknown quantity Multiplication with known quantity

5	$\left. \begin{array}{l} x + y = 7 \\ xy = 10 \end{array} \right\}$	2	1	Integer	Addition with unknown quantity Multiplication with unknown quantity
6	$\left. \begin{array}{l} x + y = 3.25 \\ 1.2x + 0.9y = 3.6 \end{array} \right\}$	2	1	Integer / Decimal	Addition with unknown quantity Multiplication with known quantity
7	$x^2 = 16$	1	1	Integer	Multiplication with unknown quantity

Note: Shaded equations were carried over from the earlier study

Table 3 lists the specific meanings proposed for each unknown in questionnaire 2 and the semantic additive and multiplicative structures inferred by such meanings. These meanings were used in pursuit of student familiarity with the exercise, for they were similar to the ones found in the textbooks reviewed.

Table 3: Meanings for unknowns in questionnaire 2

#	Equation	Meaning of unknowns	Semantic structure inferred by the proposed meaning	
			Additive	Multiplicative
1	$10,5x + 2 = 12,5$	x: number of hours needed by a plumber to complete a task		Simple proportionality
2	$x + 10 = 6x$	x: Álvaro's present age	Change or comparison	Comparison
3	$x(x + 1) = 20$	x: length of side of a rectangle	Comparison	Cartesian product
4	$\left. \begin{array}{l} 5x + 2y = 290 \\ 2x + 5y = 200 \end{array} \right\}$	x: number of cardboard boxes y: number of plastic boxes	Combination	Simple proportionality
5	$\left. \begin{array}{l} x + y = 7 \\ xy = 10 \end{array} \right\}$	x: width of floor in a rectangular room y: length of floor in a rectangular room	Combination	Cartesian product
6	$\left. \begin{array}{l} x + y = 3,25 \\ 1,2x + 0,9y = 3,6 \end{array} \right\}$	x: kilogrammes of bananas y: kilogrammes of onions	Combination	Simple proportionality
7	$x^2 = 16$	x: length of side of a square mirror	--	Cartesian product

Both groups answered questionnaires 1 and 2 in separate sessions on two consecutive days. One of the researchers, both groups' official mathematics teacher, was present. Students solved the problems individually with pencil and paper on the grounds of the following instructions provided in writing and read aloud by the teacher:

“Use each space to pose a problem from everyday life that can be solved with the equation or system of equations provided. You’ve worked with this sort of word problems in the classroom and others involving only numbers and their relationships, such as: ‘Twice the value of a number minus one is nine: figure out what that number is’. Here we’re asking you to pose problems that can arise in everyday situations. **YOU DON’T HAVE TO SOLVE THE PROBLEMS.** If you don’t know what to answer in one, you can skip it and come back to it later, after posing problems for the other equations. Please work individually and silently. If you have any questions, raise your hand and I’ll help you.”

Students were asked to pose problems from everyday situations to encourage them to attribute meaning to the unknowns and additive and multiplicative structures. The students posed no significant doubts during this data collection.

Data Analysis

The problems² posed by students were analysed first by translating them to algebraic symbolism, referred to here as ‘symbolic translation’. That involved proceeding from left to right and word for word wherever possible. Where it was not, the word problem was translated to algebraic symbolism by building a mental scheme of the mathematical relationships described.

A problem was regarded as ‘correct’ if its symbolic translation concurred with the initial equation or system of equations and ‘incorrect’ otherwise. Problems in which the symbolic translation was equivalent to the initial equation were regarded as incorrect, for the aim was to assess the conceptual understanding not of the equations as a whole, but rather of each component.

Two types of categories were defined: ‘syntactic’ and ‘semantic’, depending on whether they referred to the form or the meaning of the problems posed. The syntactic categories (Table 4) were the outcome of identifying the elements that differed between the initial equations and the symbolic translations of the problems posed by students. These categories served as the basis for meeting the first specific objective of this study. Their definition was inspired by but did not concur with the syntactic categories used in the earlier study, for the latter were refined to establish more precisely how students’ word problems diverged from the initial equations. The semantic characteristics (Table 6), which did concur with those in the earlier study, were designed to meet the second specific objective.

Table 4: Syntactic categories

Category	Name	Definition
A	Operating unknowns	In the symbolic translation of the problem posed, the unknowns operated with other elements of the algebraic expression.
B	Coefficient-unknown relationships	In the symbolic translation of the problem, the coefficients of the unknowns present were the same as in the initial algebraic expression and operated with such unknowns.
C	Number of unknowns	The number of unknowns operating with other elements in the symbolic translation was the same as in the initial algebraic expression.

² The word ‘problem’ is used here to mean word problems posed by students, irrespective of whether they meet certain minimum requirements to be regarded as problems further to a pre-established definition.

D	Terms with unknowns	The unknown was found in the same number of terms in the symbolic translation of the problem as in the initial algebraic expression.
E	Structural elements	Brackets were not added or deleted (equation 3) nor were the terms transposed in the symbolic translation of the problem.
F	Polynomial algebraic expression with equal sign	The symbolic translation of the problem yielded a polynomial algebraic expression with an equal sign.

The two possible values for category A for the word problems posed by students were ‘yes’ and ‘no’. In the first example given in Table 5, in the symbolic translation of the problem posed by a student the unknown did not operate with any other element of the equation, but rather was isolated on one side of the equal sign. Category A was consequently coded as ‘no’ in this case. The other syntactic categories included a third value, ‘not analysable (N/A)’, inasmuch as a ‘no’ in category A would preclude analysis in the rest of the syntactic categories (see example 1 in Table 5). The remaining syntactic categories were not mutually exclusive, as shown in examples 2, 3, 4 and 5 in Table 5. Category E was coded as ‘no’ whether or not the brackets were removed correctly (see example 5 in Table 5).

Table 5: Examples of syntactic categories

#	Initial equation	Word problem posed	Symbolic translation	Code						
				A	B	C	D	E	F	
1	$x^2 = 16$	Pedro has two sons. The younger, Marcos, is 4 years old. The older is twice the age of the younger. How old is the older?	$4 \cdot 2 = x$	no	n/a	n/a	n/a	n/a	n/a	n/a
2	$\begin{cases} 5x + 2y = 290 \\ 2x + 5y = 200 \end{cases}$	If I have 5 tonnes of cardboard boxes and 2 tonnes of plastic boxes, I have 290 boxes. If instead I had 200 boxes, how many boxes of each type would I have?	$\begin{cases} \frac{5}{x} + \frac{2}{y} = 290 \\ \frac{2}{x} + \frac{5}{y} = 200 \end{cases}$	yes	yes	yes	yes	yes	no	no
3	$\begin{cases} x + y = 3.25 \\ 1.2x + 0.9y = 3 \end{cases}$	Adding the kg of bananas and the kg of onions gives 3.25 kg and I have a bag with 1.2 kg of one and 0.9 kg of the other that adds up to 3.6. Calculate the system	$\begin{cases} x + y = 3.25 \\ 1.2 + 0.9 = 3.6 \end{cases}$	yes	no	yes	no	yes	yes	yes
4	$10.5x + 2 = 12.5$	A man pushes 10.5 kg of potatoes in a wheelbarrow along a road and finds two more potatoes before leaving the field. How many kilos of potatoes does he have?	$10.5 + 2x = y$	yes	no	no	no	yes	yes	yes

5	$x(x + 1) = 20$	I have a box with a number of rubbers and the same number plus 1 of pencil sharpeners and a total of 20 objects. How many rubbers are there?	$x + x + 1 = 20$	yes	yes	yes	yes	no	yes
---	-----------------	--	------------------	-----	-----	-----	-----	----	-----

As the syntactic categories provided information on incorrect problems only, those are the only results shown. In correct problems, all the categories would be coded as 'yes'. In incorrect problems, at least one of the syntactic categories was coded 'no', as Table 5 shows.

Two non-mutually exclusive semantic categories were defined as listed in Table 6. Both could be coded as either 'yes' or 'no'. Category G was not analysed in equation 7, which had no additive structure, whereas category H was analysed in all the equations, for they all involved a multiplicative structure.

Table 6: Semantic categories

Category	Name	Definition
G	Meaning of additive structures	The additive part of the problem exhibited at least one of the following semantic structures: change, combination, comparison or equalisation.
H	Meaning of multiplicative structures	The multiplicative part of the problem exhibited at least one of the following semantic structures: simple proportionality, comparison or Cartesian product.

In this analysis, problems that could not be translated to a symbolic equation (such as "Several pairs of cats have 16 kittens. How many pairs of cats are there?") were omitted and labelled as 'omitted problem'.

Results

This section discusses the classification of the problems posed by students as correct or incorrect and the use of the categories set out in the preceding section to code them. The largest number of analysable (not omitted) problems were posed for equations 2 and 7 in both questionnaires, whereas the largest number of omitted problems were posed for equation 3 in questionnaire 1 and equations 1 and 3 in questionnaire 2 (Table 7).

Table 7: Frequency of non-analysable problems (n=32)

Code	Equation/ System of equations						
	1	2	3	4	5	6	7
	Questionnaire 1						
Unanswered	3	1	13	4	3	10	0
Omitted	4	1	1	0	0	0	1
Total	7	2	14	4	3	10	1

	Questionnaire 2						
Unanswered	7	0	4	2	4	2	1
Omitted	1	0	4	0	1	2	0
Total	8	0	8	2	5	4	1

As Figure 1 shows, equations 2, 4 and 5 were the object of the largest number of correctly posed problems in both questionnaires, along with equation 7 in the second. In all these cases, the coefficients and independent terms were integers and equations 5 and 7 involved multiplication between two unknowns. The largest number of incorrect problems were posed for equations 1 and 7 in questionnaire 1 and 1 and 6 in questionnaire 2, all of which involved decimals as coefficients and independent terms. More problems were correctly posed in questionnaire 2 than questionnaire 1 for all the equations, with the widest variation in equations 3 (7 correct problems in questionnaire 1 and 19 in 2) and 7 (17 correct in 1 and 27 in 2).

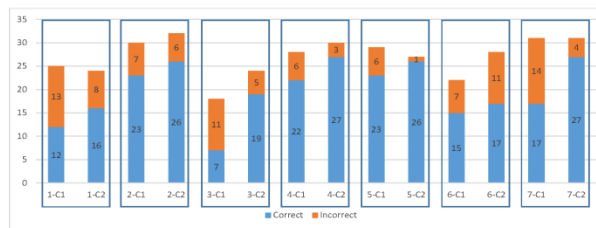


Figure 1: Number of correct and incorrect problems for questionnaires 1 and 2
Note: C1 - questionnaire 1; C2 - questionnaire 2

Syntactic categories

The syntactic categories described the differences detected between the symbolic translation of the problems incorrectly posed by students and the initial symbolic expressions. Figure 2 shows the frequencies of word problems coded as 'no' in each syntactic category. In this figure results related to each category are placed inside a rectangle. Each bar corresponds to one of the questionnaires and shows the number of problems coded as 'no' in each expressions. The divergence in students' problems from the initial equations tended to involve the relationship between coefficients and unknowns (category B), the number of unknowns defined (category C) and the number of terms in which they appeared (category D).

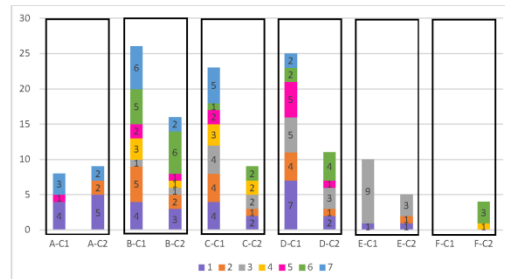


Figure 2: Frequency of word problems coded as 'no' by category
 Note: C1 - questionnaire 1; C2 - questionnaire 2

The following is a discussion of the problems posed, category by category.

Category A: Operating unknowns

In both questionnaires, the vast majority (96 %) of the (non-omitted) problems posed by students involved an unknown operating with other terms. The number of problems with no unknown operating with other members of the equation was slightly higher in questionnaire 2, specifically in connection with equations 1 and 2. Often, in the symbolic translation of the problems coded as 'no' in category A, the unknown was isolated on one side of the equal sign. That was not the case in only four omitted problems, two each on questionnaires 1 and 2, in which the answer was included in the word problem itself, obviating the unknown. One example of such instances was posed for equation 1 in questionnaire 1: "How many hours would a plumber take to finish his work if, in addition to the 10.5 hours he's already devoted, we add 2 to get 12.5?" (the symbolic translation was $10.5 + 2 = 12.5$).

Category B: Relationship between coefficients and unknowns

At least one word problem was coded no in this category in all seven equations. The equations involving decimal coefficients (equations 1 and system 6) exhibited a fairly high number of problems with a 'no' in this category in both questionnaires. Where system 6 problems were coded 'no' under this category, the relationship between coefficient and unknown was incorrect in the equation containing decimals. By way of example, one problem posed for system 6 in questionnaire 2 read as follows: "How many kg of bananas are there? and onions? If the sum of the kilogrammes of bananas and onions is 3.25 kg. In addition the shop's lorry brings 1.2 kg more bananas and 0.9 kg more onions", the total comes to 3.6 kg, and was translated symbolically as:

$$\begin{cases} x + y = 3.25 \\ x + 1.2 + y + 0.9 = 3.6 \end{cases}$$
 Note that the coefficients were not treated as such, but as independent terms.

Equation 7 stood out in this respect, with six problems in questionnaire 1 and two in questionnaire 2 in which the coefficients were not maintained. In most cases (five in questionnaire 1 and two in questionnaire 2), the difficulty was associated with the presence of exponents, for the solution to the problems posed did not involve operating with powers. In these cases students added coefficients, with the symbolic translation yielding a linear equation. The following problem from questionnaire 1 serves as an example: “A person bought four articles, but can’t remember the price of each. Calculate the price knowing she spent €16 in all and that all the articles had the same price” (symbolic translation: $4x = 16$).

Generally speaking, the number of problems in which coefficients were incorrectly related to unknowns was smaller in questionnaire 2 than in questionnaire 1. The decline was steepest in equations 2 and 7.

Category C: Number of unknowns

The number of unknowns diverged from the initial equation in at least one of the problems posed for all the equations in questionnaire 1, although it was infrequent in the systems of equations. The flaws in the word problems for equation 1 consisted in including more than one unknown or replacing one of the independent terms with an unknown quantity. In equation 2, with unknowns in two members, the five divergent problems assigned the unknowns different meanings. In equation 3, three students posed problems requiring more than one equation, apparently as a result of eliminating the brackets. In equation 7 some students cited the area of a lot but without specifying that it was square.

The number of problems with a divergent number of unknowns declined substantially in questionnaire 2 (to 9, down from the 23 in questionnaire 1). In equation 7, none of the students included a different number of unknowns than in the initial equation. In both questionnaires, divergence from the number of unknowns tended to be upward. By way of example, the following problem was posed for system of equations 4 in questionnaire 1: “I have 290 euros saved in five-euro notes and coins, but I’ve spent 90 euros. How many bills and coins do I have now?”, translated symbolically as:
$$\begin{cases} 5x + 2y = 290 \\ 5z + 2t = 200 \end{cases}$$
 The number of unknowns was reduced in only one of the problems posed for this same system of equations in questionnaire 1.

Category D: Terms with unknowns

In most cases students tended to add terms with unknowns: that happened in 20 of the 25 instances in questionnaire 1 and 6 of 10 cases in questionnaire 2. In equation 1, whenever unknowns were added (category C), the number of terms with unknowns also rose. Students also made this change in an attempt to remove the brackets in equation 3. Here and in system 5, when the multiplicative structure was replaced by an additive structure, the number of terms with unknowns was increased.

In questionnaire 1, the number of terms with unknowns was raised most frequently in equations 1, 3 and 5. The number of problems in which the

symbolic translation yielded a larger number of unknowns than the initial equation was much smaller in questionnaire 2, particularly in the aforementioned equations.

System 6 merits mention here, as it had the highest frequency of problems coded ‘no’ in questionnaire 2, where it was higher than in questionnaire 1. In this system, when the number of terms with unknowns diverged, it was smaller than in the initial system, either because a single equation was proposed or all the decimal coefficients were included as independent terms, unrelated to unknowns.

Category E: Structural elements

This category was coded ‘no’ in only a few problems. The frequency of negatives was highest for equation 3, the only one with brackets, and was lower in questionnaire 2 than in questionnaire 1. In all cases with a ‘no’ code under this category in equation 3, the problems posed by students translated symbolically to an equation without brackets. In the other word problems coded negatively in this category students transposed terms (questionnaire 1) or added brackets (questionnaire 2).

Category F: Polynomial algebraic expression with equal sign

The symbolic translation of students’ word problems failed to yield an algebraic polynomial with an equal sign in questionnaire 2 only. In the sole case involving system 4 and one of the cases involving system 6, in the symbolic translation of the students’ problem the unknown was located in the denominator and therefore did not define a polynomial. In the other two cases involving system of equations 6 the symbolic translation exhibited no equal sign.

Semantic categories

The following is a discussion of the classification of the problems into the semantic categories defined. These results provided insight into the meanings attributed by students to the operations contained in the initial symbolic expressions. The findings are listed for each equation and correct and incorrect problems are discussed jointly.

Category G: Meaning of additive structures

Further to the findings for category G (Table 8), most of the word problems posed by students attributed meaning to these structures (86 % in questionnaire 1 and 89 % in questionnaire 2)³. Notably, students found it hardest to attribute meaning to the additive structure in equation 2 in questionnaire 2, where an age context was suggested. In equation 3 students encountered difficulties even when a meaning was proposed for the unknown. In both questionnaires, the difficulties were ostensibly greater in equations than in systems.

Table 8: Word problem coding for category G

Code	Semantic structure	Equation	Total
------	--------------------	----------	-------

³ Percentages relative to analysable problems involving additive structures

		1	2	3	4	5	6	
		Questionnaire 1						
YES	Combination	16 (9)	12 (11)	6 (0)	28 (22)	27 (22)	21 (15)	110 (79)
	Change	2 (0)	12 (9)	1 (1)	0 (0)	0 (0)	0 (0)	15 (10)
	Comparison	3 (0)	0 (0)	2 (1)	0 (0)	0 (0)	0 (0)	5 (1)
	Equalisation	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)
NO		4 (3)	5 (2)	9 (5)	0 (0)	2 (1)	1 (0)	21 (11)
		Questionnaire 2						
YES	Combination	22 (16)	6 (6)	1 (0)	30 (27)	27 (26)	28 (17)	114 (92)
	Change	0 (0)	15 (13)	0 (0)	0 (0)	0 (0)	0 (0)	15 (13)
	Comparison	0 (0)	2 (0)	16 (16)	0 (0)	0 (0)	0 (0)	18 (16)
	Equalisation	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
NO		2 (0)	9 (7)	7 (3)	0 (0)	0 (0)	0 (0)	18 (10)

Note: The values in brackets denote the number of correct problems. The shaded cells indicate the semantic structure prompted by the meanings proposed for the unknowns. Equation 7 was excluded because it had no additive structure.

The occurrence of additive semantic structures in the problems posed is shown in Figure 3 for questionnaires 1 and 2. Problems involving combination prevailed in both questionnaires for both correct and incorrect answers, whereas problems involving equalisation were nearly absent and the other types exhibited a very low frequency. Combination was also observed to predominate in each algebraic expression separately, with the exception of equation 2, in which change was also frequently found in both questionnaires. The small number of problems involving comparison were proposed in the first three equations, where this category was appropriate.

More problems involving combination and comparison, most correct, were detected in questionnaire 2. The frequency of problems involving combination was particularly high in equations 4, 5 and 6 and especially low in equation 3. The meanings proposed for equations 2 and 3 in questionnaire 2 led to problems involving change or comparison, depending on the case.

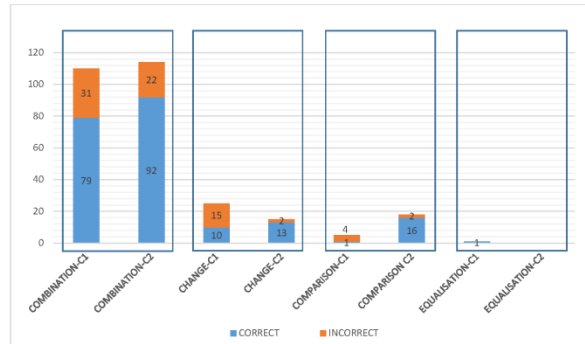


Figure 3: Frequency of additive semantic structures
 Note: C1 - questionnaire 1; C2 - questionnaire 2

Category H: Meaning of multiplicative structures

As the data in Table 9 show, most of the word problems for questionnaires 1 (62 %) and 2 (75 %) attributed meaning to multiplicative structures⁴ (category H), although the percentages were lower than for additive structures. The findings for equation 3 were particularly striking, with only one problem attributing meaning to multiplicative structures in questionnaire 1, compared to 15 in questionnaire 2. A similar difference was observed for equation 5, which also involved multiplying two unknowns.

Table 9: Word problem coding for category H

Code	Semantic structure	Equation							Total
		1	2	3	4	5	6	7	
Questionnaire 1									
YES	Comparison	1 (0)	18 (17)	0 (0)	4 (4)	0 (0)	4 (3)	3 (0)	30 (24)
	Simple proportion	13 (6)	5 (2)	1 (0)	24 (18)	2 (0)	11 (10)	4 (0)	60 (36)
	Cartesian product	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	12 (7)	13 (8)
NO		10 (5)	7 (4)	17 (7)	0 (0)	27 (23)	7 (2)	12 (10)	63 (51)

⁴ Percentages relative to analysable problems, all of which included multiplicative structures.

Questionnaire 2									
YES	Comparison	0 (0)	22 (19)	0 (0)	3 (3)	0 (0)	3 (2)	0 (0)	28 (24)
	Simple proportion	19 (15)	0 (0)	0 (0)	25 (23)	0 (0)	19 (15)	2 (0)	65 (53)
	Cartesian product	0 (0)	0 (0)	15 (14)	0 (0)	15 (15)	0 (0)	23 (22)	53 (51)
NO		5 (1)	10 (7)	9 (5)	2 (1)	12 (13)	6 (0)	6 (5)	50 (31)

Note: The values in brackets denote the number of correct problems. The shaded cells indicate the semantic structure prompted by the meanings proposed for the unknowns.

The frequency of multiplicative semantic structures for questionnaires 1 and 2 is graphed in Figure 4. Simple proportionality prevailed in both questionnaires. The difference in the number of comparison structures between the two questionnaires was nearly negligible, whilst a greater number of problems involving simple proportionality was found in questionnaire 2 (50 vs 65). The widest gap was found for the Cartesian product (13 vs 55), however, associated with the equations involving the multiplication of two unknowns (3, 5 and 7). In questionnaire 1, students proposed problems involving the Cartesian product for equation 7 only.

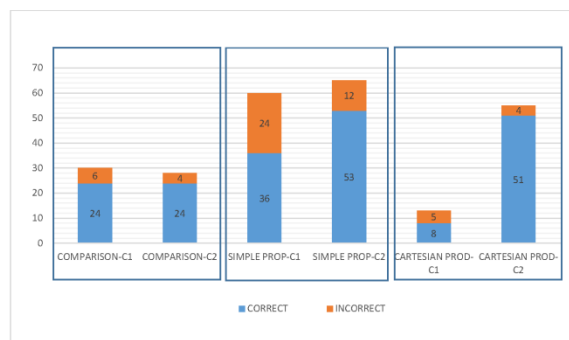


Figure 4: Frequency of multiplicative semantic structures

Note: C1 - questionnaire 1; C2 - questionnaire 2

Discussion

Analysis based on the aforementioned syntactic and semantic categories, the vehicle for meeting the two objectives of this study, provides insight into the

conceptual knowledge of algebraic symbolism acquired by compulsory secondary school students.

Although students found it difficult to pose problems that would translate in algebraic expressions concurring with the initial expressions, in most of these, over 50 % of the analysable problems posed were correct (the exceptions being 1 and 3 on questionnaire 1, where correct answers accounted for 48 % and 39 % of the totals, respectively). These difficulties eased significantly when meanings were furnished for unknowns (questionnaire 2), with over 60 % correctly worded problems in all the equations, and over 80 % in four (2, 4, 5 and 7). Students exhibited good conceptual understanding of algebraic symbolism in this second questionnaire, enabling them to attribute meaning to the equations. The presence of decimals as coefficients and equation 3, $[x(x + 1) = 20]$, with brackets and multiplication of two unknowns, revealed certain gaps in that knowledge.

The presence of decimal coefficients conditioned students' ability to pose problems. The lowest proportion of correct problems was found for equation 1 (and equation 3, analysed below) on questionnaire 1 and 1 and 6 on questionnaire 2, both bearing decimals as coefficients and independent terms. When broaching this task students failed to relate the coefficient to the unknown, tending to construe it as an independent term. Furnishing a specific meaning to the unknowns (questionnaire 2) improved performance in terms of the number of correctly worded problems for both equations, although no decline was observed in the number of problems that failed to correctly relate the coefficient to the unknown (category B). Further to that finding and given the prevalence of simple proportionality in the meanings assigned to multiplicative structures, students may be conjectured to associate multiplication with repetitive addition. That would explain why they found it hard to formulate word problems for these equations. For instance, problems such as '1.2 kilos of bananas plus 0.9 kilos of onions' or '1.2 bananas plus 0.9 onions' for the symbolic expression ' $1.2x + 0.9y$ ' in system of equations 6 are indicative of students' limited ability to put into words the precision characteristic of algebraic symbolism.

Earlier studies exploring translation in the opposite direction (from verbal to symbolic language) described several phenomena indicative of this same difficulty. Cerdán (2010) noted that students focused only on some of the words in the problem that referred to amounts, regarding amounts sharing a given word in their description to be equal. González-Calero, Arnau and Puig (2013) found students to be imprecise when specifying the meanings of letters in an algebraic expression (e.g., $x = \text{cars}$). Mitchell (2001) coined the term 'wordwalking' to mean changing the words in a problem in ways that affect their meaning, leading to interpretations that diverged from the relations described in the word problem. Rodríguez-Domingo (2015) observed that some students regarded as acceptable translations in which part of the equation was expressed more generally (e.g., an even number instead of 2). Such findings, along with the difficulties in dealing with decimals as coefficients and independent terms attested to in this study, denote a need for a sharper focus in secondary school classrooms on the importance of precision in algebraic contexts and the concomitant differences between verbal and symbolic language. Students' linguistic competence should also be developed to enable them to grasp such precision verbally.

Equation 3 stood out for its complexity, with the highest proportion of omitted problems, the smallest number of correct answers in questionnaire 1 and a substantial percentage of students who failed to attribute meaning to the additive and multiplicative structures in both questionnaires. This equation had a more complex structure than the others expressions. Divining the solution to an algebraic problem solvable with this equation would probably be more accessible to students if expressed as the system

$$\left. \begin{array}{l} y = x + 1 \\ xy = 20 \end{array} \right\}$$

and deriving equation 3 in a subsequent step in the process. Posing a problem solvable with equation 3 is cognitively demanding in terms of students' sense of structure (Vega-Castro, Molina and Castro, 2012; Hoch and Dreyfus, 2005). That notion refers to a suite of skills requiring the combined use of conceptual and procedural knowledge, including dealing with a compound term as a whole, recognising familiar structures and identifying relationships between equations or parts of an equation. To successfully pose a problem solvable with equation 3 students had to recognise x and $x+1$ as two separate unknowns, whose product is 20. That would entail understanding the expression $x+1$ as a whole. It would also involve identifying relationships between parts of the equation. The difficulties revealed in this task therefore provide insight not only into students' conceptual knowledge but also into their sense of structure.

The presence of multiplication of two unknowns had a heavy impact on students' ability to pose problems. Proof of that can be found in questionnaire 1 primarily in the number of incorrectly worded problems for equation 1, and in equation 7, where barely 50 % of the problems posed were correct. Students posed problems calling for solutions with linear equations, attesting to greater ease in attributing meaning to multiplicative structures involving coefficients and unknowns than those involving two unknowns, as observed by Fernández-Millán and Molina (2016). That difficulty was also present in terms of the meaning attributed to multiplicative structures involving unknowns. In questionnaire 1, meaning was attributed least frequently to this type of structure in equations 3, 5 and 7. Students found it easier to deal with the multiplication of two unknowns when asked to pose problems for the equations in the second questionnaire. There, where the meanings proposed for the unknowns were related to the areas of plane geometric figures, prompting the use of the Cartesian product, the number of correct problems for equations 3, 5 and 7 rose significantly.

Whole number coefficients other than 1 were also observed to render problem-posing difficult for students, corroborating findings reported by Fernández-Millán and Molina (2016). In equations 2 and 4, with coefficients other than 1, the meanings of coefficients diverged more frequently from the original than in the other equations with integers as coefficients (3 and 4). In the earlier study the authors noted that this finding might be related to a greater difficulty to verbally express multiplicative relationships with natural numbers. This circumstance was less prominent in questionnaire 2, for furnishing meaning for the unknowns helped students associate the problem with a given context.

When students experienced difficulties in posing a problem, the factors primarily affected by the divergence introduced were: the relationship between coefficients and unknowns, the number of unknowns defined and the number of terms in which they appeared. The problems posed tended to include operating unknowns. In the scant instances where that was not the case, the equations affected were the ones with the simplest structure, which were either solved or rearranged to isolate the unknown. That divergence was not observed in systems of equations, in all likelihood due to the greater complexity of the process to be followed to do so. In the earlier study, however, it was detected in the problems posed for all the equations. Other more sporadically occurring types of divergence, not detected in the earlier study, included the omission of the equal sign or of brackets, the inclusion of brackets and the positioning of unknowns in the denominator of a fraction.

Variations in the number of unknowns tended to be upward, as observed in the earlier study. Detected primarily in questionnaire 1, this divergence was attributed a number of causes depending on the type of equation. Particularly prominent were flawed verbal expression when the equations contained decimals or the product of equal unknowns, which induced students to include more than one unknown. In such cases and where multiplicative were replaced with additive structures, students tended to raise the number of terms with unknowns.

Further to the information gleaned from the semantic categories, students attributed meaning to additive structures in nearly 90 % of the problems posed, exhibiting greater uncertainty in dealing with multiplicative structures, primarily in questionnaire 1. Combination followed by change prevailed in additive structures, as in the earlier study. These two are the types of additive structures most frequently found in primary school textbooks, according to a review by Orrantia, González and Vicente (2005). The paucity of problems involving comparison or equalisation was common to this and the earlier study.

When a specific meaning was furnished for the unknowns (questionnaire 2), students attributed meaning to multiplicative with the same ease as to additive structures (146 problems vested meaning in multiplicative and 147 in additive structures). That did not translate into a significant overall rise in the number of problems attributing meaning to additive structures in questionnaire 2 relative to questionnaire 1, although some of the semantic structures were impacted: in equation 2 there were more additive problems involving change and in equation 3 more involving comparison.

Although students did not tend to pose additive problems involving comparison, when induced to do so by the meaning furnished for the unknowns, comparisons (e.g., age) were frequently used and most of the problems posed were correct. Findings on the use of the Cartesian product in multiplicative problems were analogous. That multiplicative structure was scantily present in questionnaire 1 and absent altogether in equations 3 and 5, which involved multiplying two unknowns. No more than two students assigned meaning based on that product in both. Nonetheless, when meanings associated with lengths were proposed for the unknowns, all students used the Cartesian product in both equations, and nearly all correctly.

On the whole, simple proportionality and comparison were the prevalent multiplicative semantic structures in questionnaire 1, as was also reported in Fernández-Millán and Molina (2016).

The questionnaire 1 results concurred with the earlier study in detecting specific meanings for the operational structures that were weakly associated with such operations: Cartesian product and additive comparison. Thinking of the context of area of plane rectangles, for instance, would have helped students to pose problems for equations involving the multiplication of two unknowns in questionnaire 1. In the absence of situations associated with the Cartesian product, students found it difficult to pose problems correctly in these cases. The use of simple proportionality and comparison to attribute meaning to such equations is artificial, for it entails posing a situation in which a) both the scaling factor and one of the quantities for comparison is unknown or b) the number of elements in each group and the number of groups is unknown.

Such difficulties were drastically reduced, however, when a meaning was furnished for the unknowns, suggesting that conceptual knowledge was partially connected. The results inferred that such knowledge lies in the zone of proximal development and inaccessible to students if unaided. Although the equations used were all familiar to them, students needed help to connect them to the classroom contexts to which they were accustomed.

Lastly, the order in which the equations were presented was found to be unrelated to both the number of problems posed and the number of correctly worded problems (Table 7 and Figure 1).

Conclusions

This article, the continuation of an earlier study by Fernández-Millán and Molina (2016), compares the findings from both studies while further exploring the conceptual understanding of algebraic symbolism acquired by two groups of students in the last year of compulsory secondary school. The results of this second study, which are more promising, suggest the potential for compulsory algebra instruction to develop students' conceptual knowledge, although greater attention should be paid to the semantic aspects of algebra if students are to access such knowledge unaided.

The findings gleaned from the first questionnaire used (free problem-posing) corroborated the results of the earlier study in terms of the difficulties experienced by students in posing problems for equations involving the multiplication of two unknowns and coefficients other than 1. In both studies problem-posing was particularly difficult for equation 3, which was interpreted to signify shortcomings in students' sense of structure. Limitations were also detected in students' ability to grasp the precision expressed with algebraic symbolism in verbal language. The tendency to isolate the unknown detected in the earlier study was not corroborated here, however, suggesting that students' concept of algebraic expressions was more relational than operational. In both studies, the predominant semantic categories were combination followed by change, in additive structures, and simple proportionality followed by comparison in multiplicative structures. The least prevalent additive semantic structures were comparison and equalisation and the least multiplicative structure was the Cartesian product.

The findings ratified the need to pay greater attention to expression through verbal representation of relationship schemes that can be modelled using equations and systems of equations, as well as to decimal coefficients and coefficients other than 1. The development of linguistic competence in algebraic contexts calls for steady work that can be undertaken in arithmetic contexts, given the wealth of elements and meanings shared by algebraic and numerical symbolisms.

From the educational standpoint, the shortcomings identified provide insight for the design of instructional proposals geared to developing students' understanding of the meaning of arithmetic operations and algebraic symbolism. The study ratifies the utility of problem-posing as a useful tool for evaluating student's implicit conceptual knowledge, whether for educational or research purposes.

Acknowledgements

This study, which forms part of the first author's PhD. thesis, was conducted under Spanish Research and Development Projects EDU2013-41632-P and EDU2016-75771-P, funded by the Spanish Ministry of Economy and Competitiveness.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Elena Fernández-Millán - Universidad de Granada, Spain.

Marta Molina - Universidad de Granada, Spain.

References

- Álvarez, I. & Gómez-Chacón, I. M. (2015). Understanding the algebraic variable: Comparative Study of Mexican and Spanish students. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(6), 1507-1529
- Arcavi, A. (1994). Symbol sense: informal sense-making in formal mathematics. *For the Learning of Mathematics*, 1(3), 24-35.
- Arcavi, A. (2005). Developing and using symbol sense in mathematics. *For the learning of mathematics*, 25(2), 42-47.
- Arnau, D. & Puig, L. (2013). Actuaciones de alumnos instruidos en la resolución algebraica de problemas en la hoja de cálculo y su relación con la competencia en el método cartesiano. *Enseñanza de las Ciencias*, 31(3), 49-66. DOI: <https://doi.org/10.5565/rev/ec/v31n3.967>
- Bills, L. (2001). Shifts in the Meanings of Literal Symbols. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the Twenty-fifth PME International Conference* (pp. 2-161, 2-168). Utrecht, The Netherlands: PME.
- Booth, L.R. (1984). *Algebra: Children's strategies and errors*. Windsor, UK: NFER-Nelson.
- Capraro, M. & Joffrion, H. (2006). Algebraic Equations: Can Middle-School Students Meaningfully Translate from Words to Mathematical Symbols? *Reading Psychology*, 27(2-3), 147-164.

- Carpenter, T. & Moser, J. (1982). The development of addition and subtraction problem solving skills. In T. Carpenter, J. Moser & T. Romberg (Eds.), *Addition and subtraction: Developmental perspective* (pp. 9-24). Hillsdale, N. Jersey: Lawrence Erlbaum Associates.
- Castro, E. (2001). Multiplicación y división. In E. Castro (Ed.), *Didáctica de la matemática en educación primaria*. Madrid, España: Síntesis, pp. 203-230.
- Castro, E. & Castro, E. (1997). Representaciones y modelización. In L. Rico (Ed.), *La educación matemática en la enseñanza secundaria* (pp. 95-124). Barcelona, España: Horsori.
- Castro, A., Prat, M. & Gorgorió, N. (2016). Conocimiento conceptual y procedimental en matemáticas: su evolución tras décadas de investigación. *Revista de Educación*, 374, 43-68
- Cedillo, T. E. (2001). Toward an algebra acquisition support system: A study based on using graphic calculators in the classroom. *Mathematical Thinking and Learning*, 3, 221–259.
- Cerdán, F. (2010). Las igualdades incorrectas producidas en el proceso de traducción algebraico: un catálogo de errores. *PNA*, 4(3), 99-110.
- Chalouh, L. & Herscovics, N. (1988). Teaching algebraic expressions in a meaningful way. In A.F. Coxford & A.P. Shulte (Eds.), *The ideas of algebra K-12* (pp. 33-42). Reston, Virginia: NCTM.
- Chappell, M. F. (2001). Creating connections: Promoting algebraic thinking with concrete models. *Mathematics Teaching in the Middle School*, 7, 20–25.
- Crooks, N. & Alibali, M. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, 34(4), 344-377
- Fernández-Millán E. y Molina, M. (2016). Indagación en el conocimiento conceptual del simbolismo algebraico de estudiantes de secundaria mediante la invención de problemas. *Enseñanza de las Ciencias*, 34(1), 53-71. DOI: <http://dx.doi.org/10.5565/rev/ensciencias.1455>
- Ferrucci, B. J., Kaur, B., Carter, J. A. & Yeap, B. (2008). Using a model approach to enhance algebraic thinking in the elementary school mathematics classroom. In C. E. Greenes & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics: 2008 Yearbook* (pp. 195–209). Reston, VA: National Council of Teachers of Mathematics.
- Filloy, E. & Rojano, T. (1989). Solving equations: The transition from arithmetic to algebra. *For the learning of Mathematics*, 9(2), 19-25.
- Filloy, E., Rojano T. & Puig, L. (2008). *Educational algebra. A theoretical and empirical approach*. New York, NY: Springer.
- Fujii, T. & Stephens, M. (2008). Using number sentences to introduce the idea of variable. In C. E. Greenes & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics: 2008 Yearbook* (pp. 127–140). Reston, VA: National Council of Teachers of Mathematics.
- Furinghetti, F. & Paola, D. (1994). Parameters, unknowns and variables: a little difference? In J.P. da Ponte & J.F. Matos (Eds.), *Proceedings of the XVIII International Conference for the Psychology of Mathematics* (vol. 2, pp. 368-375). Lisboa, Portugal: Universidad de Lisboa.
- Gómez, P. (2007). *Desarrollo del conocimiento didáctico en un plan de formación inicial de profesores de matemáticas de secundaria*. PhD Thesis. Granada: Universidad de Granada.
- González-Calero, J. A., Arnau, D. & Puig, L. (2013). Dificultades en la construcción de nombres de cantidades durante la resolución algebraica de problemas verbales por estudiantes de primaria. In A. Berciano, G. Gutiérrez, A. Estepa, & N. Climent (Eds.), *Investigación en Educación Matemática XVII* (pp. 301-310). Bilbao, Spain: Sociedad Española de Investigación en Educación Matemática.
- Herscovics, N. & Kieran, C. (1980). Constructing meaning for the concept of equation. *Mathematics Teacher*, 73(8), pp. 572-580.
- Hiebert, J. & 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.
- Hoch, M. & Dreyfus, T. (2005). Students' difficulties with applying a familiar formula in an unfamiliar context. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 29th conference of*

- the international group for the psychology of mathematics education* (Vol. 3, pp. 145–152). Melbourne, Australia: University of Melbourne.
- Kaput, J. (1989). Linking representations in the symbolic systems of algebra. In S. Wagner & C. Kieran (Eds.), *Research agenda for mathematics education: Research issues in the learning and teaching of algebra* (pp. 167-194). Reston, VA: NCTM.
- Küchemann, D.E. (1981). Algebra. In K.M. Hart, M.L. Brown, D.E. Küchemann, D. Kerslake, G. Ruddock & M. McCartney (Eds.), *Children's understanding of mathematics: 11-16* (pp. 102-119). London, Reino Unido: John Murray,
- Lin, P. J. (2004). Supporting teachers on designing problem-posing tasks as a tool of assessment to understand students' mathematical learning. In M. Hoines y A. Fuglestad (Eds.), *Proceedings of the 28th annual meeting of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 257-264). Bergen, Noruega: Bergen University College.
- Mestre J. P. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Journal of Applied Developmental Psychology*, 23(1), 9-50.
- Mitchell, J. M. (2001). Interactions between natural language and mathematical structures: the case of "wordwalking". *Mathematical Thinking and Learning*, 3(1), 29-52.
- Molina, M. (2014). Traducción del simbolismo algebraico al lenguaje verbal: indagando en la comprensión de estudiantes de diferentes niveles educativos. *La Gaceta de la RSME*, 17(3), 559-579.
- Molina, M., Rodríguez-Domingo, S., Cañadas, M.C. & Castro, E. (2016). Secondary School Student's Errors in the Translation of Algebraic Statements. *International Journal of Science and Mathematics Education*, 13(3), 1-20.
- Ng, S. F. & Lee, K. (2009). The model method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education*, 40(3), 282–313.
- Orrantia, J., González, L.B. & Vicente, S. (2005). Un análisis de los problemas aritméticos en los libros de texto de Educación Primaria. *Infancia y Aprendizaje*, 28, 420-451. <http://dx.doi.org/10.1174/021037005774518929>
- Resnick, L.B., Cauzinille-Marmeche, E. & Mathieu, J. (1987). Understanding algebra. In J.A. Sloboda & D. Rogers (Eds.), *Cognitive process in mathematics* (pp. 169-203). Oxford, Reino Unido: Clarendon Press.
- Rittle-Johnson, B. & Schneider, M. (2015). Developing Conceptual and Procedural Knowledge of Mathematics. In R.C. Kadosh & A. Dowker (Eds.), *Oxford Handbook of Numerical Cognition* (pp. 1118-1134). Oxford, UK: Oxford University Press
- Rittle-Johnson, B. & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, 99, 561–574. doi: 10.1037/0022–0663.99.3.561.
- Rodríguez-Domingo, S. (2015). *Traducción entre los sistemas de representación simbólico y verbal: un estudio con alumnado que inicia su formación algebraica en secundaria*. PhD thesis. Granada: Universidad de Granada.
- Ross, A. & Willson, V. (2012). The Effects of Representations, Constructivist Approaches, and Engagement on Middle School Students' Algebraic Procedure and Conceptual Understanding. *School, Science and Mathematics*, 112(2), 117-128.
- Sheikhzade, M. (2008, July). *Promoting skills of problem-posing and problem- solving in making a creative social studies classroom*. Presented at 4th Global Conference, Oxford. Available at <http://www.inter-disciplinary.net/ati/education/cp/ce4/Sheikhzade%20paper.pdf>.
- Stoyanova, E. & Ellerton, N. F. (1996). A framework for research into students' problem posing in school mathematics. In P. C. Clarkson (Ed.), *Technology in mathematics education (Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 518–525). Melbourne: Mathematics Education Research Group of Australasia.

- Usiskin Z. (1988). Conceptions of School Algebra and Uses of Variables. In Coxford A.F. & Shulte A. P. (Eds.), *The Ideas of Algebra K-12* (pp. 8-19). VA: National Council of Teachers of Mathematics.
- Vega-Castro, D., Molina, M. & Castro, E. (2012). Sentido estructural de estudiantes de bachillerato en tareas de simplificación de fracciones algebraicas que involucran igualdades notables. *Relime*, 15(2), 233–258.
- Wagner, S. (1981). An analytical framework for mathematical variables. In Equipe de Recherche Pedagogique (Eds.), *Proceedings of the 5th International PME Conference* (pp.165-170). Grenoble, France: PME.