A Systematic Analysis of Errors in the Simplification of a Rational Expression

K.A. Ruhl[a], S.R. Belward[a] and J. Balatti[b]
[a]*School of Engineering and Physical Sciences, James Cook University, Townsville, Australia;*
[b]*School of Education, James Cook University, Townsville, Australia*
karen.ruhl@my.jcu.edu.au

Exploring the errors that mathematics students frequently make is a means by which teachers can gain a better understanding of students' difficulties. Reported here are the process by which the algebraic working of 95 undergraduate students who incorrectly simplified a rational expression was analysed and the results of the analysis. Initially, a deductive approach to analysing the errors was planned, categorising students' mistakes using the error types identified, named and described in the literature. In reviewing the literature, however, it became clear that this would be no simple task. The large body of literature, while rich in examples of "typical errors" that could be expected in students' working, had two limitations. Firstly, the error types lacked precise descriptions and were mainly described by example only. Secondly, insufficient details of the procedures used to categorise the errors prevented replication of the categorising process. Consequently, a mainly inductive approach, that categorised the errors by their location and inferred student operation was devised. This systematic approach resulted in generating descriptions of three error categories.

Introduction

Exploring the errors that mathematics students frequently make helps teachers gain a better understanding of students' difficulties. In articulating the benefits of approaching errors as a positive part of the learning process, Ashlock [1] writes (p.9): "As we teach computation procedures, we need to remember that our students are not necessarily learning what we think we are teaching; we need to keep our eyes and ears open to find out what our students are actually learning. We need to be alert for error patterns!" An important first step in analysing errors that students make is categorising those errors. Error categories help illustrate the patterns in the way students make errors. Furthermore, categorizing errors allows teachers and researchers to explore teaching strategies that may prevent or remedy errors that lie within the whole category.

In this paper we report on the findings of a study that investigated the errors undergraduate students made in simplifying a rational expression. The researchers anticipated using existing literature to categorise the errors observed. However, an extensive literature review found error categories that lacked the precision required for certainty over the exact nature of the errors located within each category. Moreover, in attempting to generate their own categories, the researchers found a lack of literature presenting detailed procedures for the development of error categories. Therefore in addition to identifying the errors students make in simplifying a rational expression, this paper has an additional aim of demonstrating a transparent and easily replicated procedure for the generation of error categories.

The remainder of the paper is organised as follows. The contribution of the work reported here to the field is established through a review of literature on existing methods for categorising errors, predominantly focussing on errors made in the simplification of a rational expression. Following this, we present a mainly inductive process to categorising the errors made by 95 tertiary preparatory mathematics students when simplifying a rational expression question. This approach resulted in three distinct main error categories. Finally we discuss the results in the context of the existing literature and propose areas for further work.

Literature Review

The literature reveals there are varying approaches to the categorising of errors students make when performing algebra. In general, a common approach to collecting data is to use tests [2-5]. Other approaches include student interviews, student reflections [6] and observations [7]. Sometimes more than one type of data is collected.

Of the literature reviewed, it would appear that approaches to error categorising in algebra are dominated by an inductive approach [2, 5, 6]. For example Storer [2] analysed incorrect solutions to a 52 question test on algebraic fractions to produce 15 error categories, each labelled with a theme that describes features or inferred causes of the errors. Poon and Leung [5] presented twenty-one error categories resulting from the statistical analysis of students' work in an algebra test as well as input from selected teachers. Again, the categories consisted of themes that describe features or inferred causes of the errors in the incorrect solutions analysed. One exception to the inductive approach is the work of Payne and Squibb [4] that used 23 pre-existing error categories from Sleeman [3] and three pre-existing error categories from Matz [8], before defining new error categories to take account of the remaining data collected in their study. Payne and Squibb [4], derived error categories by examining student working on a test of 56 questions, all requiring the solution of a linear equation in one unknown. Error categories were defined with as few as a single occurrence from the working of 86 students across all questions. This is in stark contrast to the pre-existing error categories they used from the work of Sleeman [3] which were defined only if consistent behaviour leading to the same incorrect operations were observed across a sequence of similar questions.

In general, the most common form of output from error analysis is a list of categories, frequencies and examples illustrating each category (see for example, [2-5]). This kind of research informs teaching practice and furthers understanding of student thinking. The study of Storer [2] cites discussions on teaching practice as a motivator for the categorising of errors. The list of errors produced, along with an analysis of their frequency in the solutions students provided was intended to contribute to such discussions. Some studies explicitly focus attention on the use of error categories for understanding student thinking. In these cases, the lists of categories are often of secondary concern and causal links explaining the thinking leading to the observed errors are proposed [3, 4 ,6 ,7]. Additional data are sometimes used to investigate the causal link, for example Sleeman [3] used interviews. In other cases, the researcher may infer the causes for the errors [7].

Two common weaknesses emerge from reviewing the literature on error categories. Firstly, the process through which the error categories arise is often unclear. Error categories are presented without a detailed procedure that allows subsequent researchers to replicate the study. For example in the work of Storer [2] and Poon and Leung [5], there is no detail regarding how the error categories were constructed, beyond the implied thinking of the researchers that the labeling of the categories suggests. Other studies, such as the work undertaken by Carry, Lewis and Bernard [6], appear overly complicated making them difficult to understand or replicate. One exception to this is the work of Sleeman [3], where it is possible to reconstruct the study and arrive at a set of error categories that is precisely described.

Secondly, there is often a lack of clarity with how error categories are described. This is caused by authors providing short descriptions of errors usually with some examples, rather than precise definitions (see for example, Storer [2] and Poon and Leung [5]). This lack of precision perhaps explains researchers' lack of referencing to previously defined errors. Lack of precision may also lead to ambiguity.

Such ambiguity is present in literature that categorises errors in the simplification of rational expressions or solution of equations involving rational expressions. The term "cancellation" is frequently used as a descriptor for these errors. Matz [7] describes "cancellation errors" as having the form $\frac{AX+BY}{X+Y} \Rightarrow A+B$. She states (p. 118) for some problems, "the way partial answers are composed into a final answer (superficially) appears more ad hoc. Sometimes signs (particularly minus signs) are ignored, slashed out literals are variously treated as 0 or 1, and not all partial answers always figure in the final answer". Barnard [8] describes two quite different examples, $\frac{x+y}{y} \Rightarrow x$ and $\frac{x+y}{y} \Rightarrow x+1$ as "inappropriate cancelling". Similarly, Parish and Ludwig [9] list examples of errors that are described as "cancellation", such as when $\frac{x+2}{2} = 3$ is incorrectly simplified to $x = 3$. These authors note that they prefer to call this process "obliteration". Poon and Leung [5] categorise the error $\frac{4-12y}{4} \Rightarrow 1-12y$ as: "Misunderstand[ing] the operation of algebraic fractions".

In all of these examples, it appears likely that some form of cancellation occurs; however, the exact mechanics of the cancellation process is unclear. In some cases the most likely inference is that cancellation involves forming a quotient from parts of an expression, while in others, it appears as though the cancellation involves subtraction between quantities on the numerator and the denominator. With the exception of Matz [7], the error descriptions in the literature cited above do not adequately provide other researchers with error categories errors that are unambiguous. We conclude that desirable error categorising protocols should provide categories that have precise definitions and that the method by which the categories have been determined should be detailed enough to be replicated by other researchers.

Method

The research reported here is part of a larger study [10] investigating student learning, in particular, student errors, in the algebraic component of an undergraduate preparatory mathematics course at an Australian university. The one semester course is equivalent to the secondary school mathematics course that prepares students for entry into disciplines such as engineering or the natural sciences where knowledge of calculus is required. A range of students enrol in the course; some have not satisfied mathematics prerequisites for entry to the degree of their choice, while others are enrolled in degree programs that have no mathematics prerequisite for entry but are required to study this level of mathematics during their degree. It is assumed that students enrolled in this course do not have any prior algebraic knowledge.

Data collection instrument

Students sat for the algebra test of 20 questions after having completed the five week long algebra component which comprised approximately the middle third of the course. The test, taken under formal exam conditions, was worth 15% of the total assessment. Students were directed to show all their working for each question attempted. They were also asked to indicate their level of confidence on a five point scale for each question. In compliance with ethics requirements, the analysis of the data took place in the semester following the delivery of the course. Of the 160 students enrolled in the course in 2010, 151 students had volunteered for the study and of these, 133 sat the test.

The analysis reported here is of the incorrect solutions to the question requiring the simplification of a rational expression. The task required students to

Simplify the following rational expression completely:

$$\frac{b^3 + 6b}{3b}$$

The standard form solution the students had been taught was to factorise the numerator and then cancel common factors from the numerator and denominator:

$$\frac{b^3 + 6b}{3b} = \frac{b(b^2 + 6)}{3b} = \frac{\not{b}(b^2 + 6)}{3\not{b}} = \frac{b^2 + 6}{3}$$

For this question, 113 of the 133 students (85%) provided a solution with at least one error.

Confidence/memory indicator

To ensure the quality of the data, these113 responses were filtered using student confidence ratings in an attempt to remove solutions that involved guesswork. The responses from students who used the confidence/memory indicator to indicate that they had "forgotten how to do this type of question altogether" (10) or "didn't remember seeing this type of question before" (2), were removed. It was assumed that these responses would contain a significant level of guesswork. Similarly, the responses of students who had left the confidence/memory indicator blank (6) were removed, as it was not possible to determine if their working involved guesswork. Using this filtering process, 18 of the 113 responses were excluded from the data set. This left 95 incorrect solutions for coding in which students had indicated that they had been "confident that they were right" (16), "fairly confident that they were right" (49) or "had forgotten how to do <u>bits</u> of this type of question" (30).

Procedure used for coding students' algebraic working

The process used to analyse the solutions produced a hierarchical coding structure with the "core codes", at the top level of the hierarchy. Stepwise, the procedure was as follows:

Step 1: The starting point in coding each response was to identify the first process the student appeared to use in their solution. This first step produced two core codes, namely, *Attempted to simplify without factorising* (80), *Attempted to factorise* (15). The responses in each of these core codes were then coded further.

Step 2: The second and subsequent steps in the analytic process involved systematic coding of the working for the solutions in each of the core codes. The coding process involved coding every element of the solution in the order in which it appeared until all the working in the solution was exhausted. An "element" may refer to an operator, a term or a factor of a term. The process involved producing error descriptions in terms of the following three dimensions:

a) Each element was coded as *correct* or *incorrect* within the context of the prior working relating to that element.
b) When it was incorrect, inferences were made and recorded concerning the

location from which the error appeared to have emanated.

c) Each error was also described in terms of the *operations* that the student appeared to have taken to arrive at that particular error.

Step 3: The third step produced the error categories. Using content analysis, the error descriptions were further coded, with a focus on the operation that the student appeared to have performed that led to the error. Error categories emerged from grouping the errors which demonstrated similar operations. Where the data yielded 10 or more individual occurrences of an error, the researchers named the error category and developed a detailed description of it.

Results

The output from Step 1 and Step 2 is presented in Tables 1-4 (see Appendix). These steps led to a total of 26 error descriptions. Common amongst the errors recorded are the mathematical operations of ratio, difference, ratio of like terms but retaining the variable[6] and factorising. Also evident were errors appearing to involve the interpretation of a sum of two terms as their product and the interpretation of an index as a coefficient.

The grouping of the errors performed in Step 3 is shown in Table 5. To illustrate this process note that the 48 pieces of student working that is recorded in the fourth column of Table 5 is composed of the errors from the second, fourth and fifth columns of Table 1, the second and eighth columns of Table 2 and the third, ninth and tenth columns of Table 4. All of these errors involved forming a quotient in error.

The grouping process led to the identification of six different error categories. In Table 5 these errors are briefly described in the fifth column and named in the sixth column. The description of the three main error categories (where the data yielded 10 or more individual occurrences of an error) is presented in detail below:

The *simple cancellation error* occurs when a rational expression is incorrectly simplified by cancelling a factor that is common to at least one of the terms on each of the numerator and denominator, but that is also not common to all terms on the numerator and denominator.

The *cancellation by subtraction error* occurs when a rational expression is incorrectly simplified in such a way that the resulting expression appears to involve the difference between like terms or coefficients on the numerator and the denominator.

The *cancellation by division of coefficients retaining the variable error* occurs when a rational expression is incorrectly simplified using the ratio of coefficients of two like terms, one on the numerator and one on the denominator where the resultant term is the ratio of the coefficients multiplied by the common variable.

The most commonly observed error was the *simple cancellation* error, (with 48 errors, or 47% of all inferable errors being coded in this category). More than twice the number of errors was recorded in this category than in any other individual category. The other two most frequently recorded errors were the cancellation by subtraction error (22, 21%) and the cancellation by division of coefficients retaining the variable error (20, 18%).

---

[6] An example of *ratio of like terms retaining the variable* is so the quotient is processed properly for the coefficients, but not the variable.

Table 5. Coding of Errors Into Categories

Table and location within	Error description	# of errors	Total	Brief error description	Error Name
Table 1, column 2, row 5	The b^2 terms appears to result from $\frac{b^3}{b}$. The $2b$ term appears to result from $\frac{6b}{3}$.	2	48	Resultant term appears as a **ratio** of circled components.	Simple cancellation error
Table 1, column 4, row 5	The integer 2 appears to result from $\frac{6b}{3b}$.	20			
Table 1, column 5, row 5	The coefficient of $2b$ appears to result from $\frac{6}{3}$.	15			
Table 2, column 2, row 6	The integer 2 appears to result from $\frac{2b}{b}$.	3			
Table 2, column 8, row 6	The b^2 term appears to result from $\frac{b^3}{b}$.	1			
Table 4, column 3, row 17	The term $2b$ appears to result from $\frac{6b}{3}$.	1			
Table 4, column 9, row 11	The integer 2 appears to result from $\frac{6}{3}$.	1			
Table 4, column 10, row 17	The integer 2 appears to result from $\frac{6}{3}$.	5			
Table 1, column 6, row 5	The $3b$ appears to result from $6b - 3b$.	10	22	The resultant term or coefficient appears as the **difference** between circled components.	Cancellation by subtraction error
Table 1, column 7, row 5	The coefficient of the $3b$ appears to result from $6 - 3$.	3			
Table 2, column 3, row 6	The term b appears to result from $2b - b$.	3			
Table 2, column 7, row 6	The $2b$ term appears to result from $3b - b$.	1			
Table 2, column 9, row 12	The $5b$ appears to result from $6b - b$.	1			
Table 4, column 6, row 11	The $3b$ appears to result from $6b - 3b$.	1			
Table 4, column 8, row 11	The $2b$ term appears to result from $3b - b$.	2			
Table 4, column 11, row 17	The integer 3 appears to result from $6 - 3$	1			
Table 1, column 3, row 5	The coefficient of the $2b$ term appears to result from $\frac{6}{3}$. The variable b is retained	19	20	Resultant term appears with a coefficient that is a **ratio** of the circled coefficients, while the variable b is **retained**.	Cancellation by division of coefficients while retaining the variable error
Table 3, column 6, row 6	The coefficient of the $3b$ term appears to result from $\frac{9}{3}$. The variable b is retained	1			
Table 1, column 8, row 5	The $9b$ appears to result from $3b + 6b$.	2	3	The resultant numerator appears as if the b^3 has **changed** form to $3b$ then added to another term.	Changing form
Table 2, column 9, row 6	The numerator appears to result from $3b + 3b$.	1			
Table 1, column 9, row 5	The $6b^4$ appears to result from $b^3 \times 6b$.	1	6	The resultant term appears as a **product** of the circled components	Conjoining
Table 2, column 5, row 6	The $2b^4$ appears to result from $b^3 \times 2b$.	2			
Table 2, column 3, row 12	The b^4 appears to result from $b^3 \times b$.	1			
Table 2, column 10, row 6	The result appears as $2 \times \frac{b^3}{3b}$. Their result may have been treated as a mixed fraction.	1			
Table 3, column 3, row 6	The $2b^3$ appears to result from $b^3 \times 2$.	1			
Table 4, column 2-5, row 5	One term of the numerator is factorised incorrectly	4	4	Only one term is factorised correctly.	Factorising error

Discussion and Conclusions

Overall, students made one of two key choices (the core codes) in attempting to answer the question in this study. Most incorrect solutions (over 80%) resulted from choosing to simplify the rational expression without appreciating that such simplification entailed identifying and then cancelling factors common to all terms in the numerator and denominator.

The approach used to categorise the errors generated descriptions of the algebraic processes that appeared to result in the errors. These were then clustered into categories according to the main operation evident in the error descriptions. The error categories generated give a detailed picture of the difficulties students experienced when simplifying this rational expression.. Valuable information would have been lost if the student working had been categorised using the more general error categories outlined in the literature review. For example, these students' errors could have been categorised using Poon and Leong's [5] "Misunderstand[ing] the operation of algebraic fractions", Barnard's [8] "inappropriate cancelling" or Matz's [7] description of a "cancellation error". In so doing, only a limited indication of the underlying mechanism behind the error would have been apparent.

In this study, the most common underlying mechanisms behind the errors involved incorrect ratios; incorrect differences; and forming ratios of like terms, but retaining the variable. This resulted in the three main error categories presented in the previous section.

Errors similar to the *simple cancellation error* and the *cancellation by subtraction error* can be found in the literature. For example, Carry, Lewis and Bernard [6] describe similar types of errors as "operator errors". In a list of 37 examples (page 43-45), they describe these errors as "a collection of errors which involve the deletion of elements from expressions". The authors explain that these errors appear as approximations of the valid operations of "subtraction from both sides of an equation, division of both sides, division of quotients and subtraction". Here the authors do subcategorise these errors. Of particular interest, 22 of these errors are subcategorised as errors in the "simplification of quotients", while another three are subcategorised as involving "subtracting s[a sub expression] from terms containing it". At first glance the names of these subcategories suggest possible correspondence with the categories *simple cancellation error* and *cancellation by subtraction error* that have been defined in this study. However, on closer inspection it is observed that there is considerable breadth in the categories defined by Carry, Lewis and Bernard [6]. This is almost certainly a consequence of the breadth of their study. The trade-off between depth and breadth in the definitions of error categories needs careful thought, and needs to be considered in the context of the intended outcomes of the error analysis.

The researchers are unaware of any publication describing an error of the form of *cancellation by division of coefficients retaining the variable* error. This is an area for further research. In particular it would useful to see if this result is repeated in similar questions with other cohorts and if so, to investigate student thinking leading to this error.

In conclusion, the analysis of the errors across the whole data set produced three error categories that were defined precisely. The results provide teachers with a snapshot of the difficulties students have in simplifying a rational expression of the type selected. The findings, however, are limited to the simplification of only one example of a rational expression by only one cohort of tertiary preparatory mathematics students. Further research is required with other forms of rational expressions and with more students. More research is also required to test the applicability of the error analysis process documented and illustrated here to the analysis of student errors in solutions to a broader range of problems.

References

[1] R.B. Ashlock, Error Patterns in Computation: Using Error Patterns to Improve Instruction. 2006, Upper Saddle River: Pearson Prentice Hall.

[2] W. Storer, An analysis of errors appearing in a test on algebraic fractions. The Mathematical Gazette. 40(331) (1956), pp. 24-33.

[3] D. Sleeman, An attempt to understand students understanding of basic algebra. Cognitive Science: A Multidisciplinary Journal. 8(4) (1984), pp. 387-412.

[4] S. Payne and H. Squibb, Algebra mal-rules and cognitive accounts of error. Cognitive Science: A Multidisciplinary Journal. 14(3) (1990), pp. 445-481.

[5] K. Poon and C. Leung, Pilot study on algebra learning among junior secondary students. International Journal of Mathematical Education in Science and Technology. 41(1) (2010), pp. 49-62.

[6] L.R. Carry, C. Lewis, and J. Bernard, Psychology of Equation Solving: An Information Processing Study. Final Technical Report. Department of Curriculum and Instruction, University of Texas at Austin, Austin, USA, 1979.

[7] M. Matz, Towards a computational theory of algebraic competence, Journal of Mathematical Behavior. 3 (1980), pp. 93-166.

[8] T. Barnard. Hurdles and strategies in the teaching of algebra: Part II, Mathematics in School. 31(4) (2002), pp. 41-44.

[9] C. Parish and H. Ludwig, Language, Intellectual Structures and Common Mathematical Errors: A Call for Research, School Science and Mathematics. 94(5) (1994), pp. 235-39.

[10] K. A. Ruhl, Aspects of the tertiary preparatory mathematics students' competence and confidence with algebra: A case study through the lens of error analysis, BSc (Hons), James Cook University,Townsville, 2011.