International Journal of Engineering Education Vol. 33, No. 2(A), pp. 662–671, 2017
Printed in Great Britain
0949-149X/91 $3.00+0.00

Analysing Errors Students Make in Summative Tests: A Case Study in Research-Led Engineering Education*

L. C. WOOLLACOTT and L. VAN DYK
School of Chemical and Metallurgical Engineering, University of the Witwatersrand, South Africa.
1 Jan Smuts Av, Braamfontein, Johannesburg, South Africa, 2050. E-mail: lorenzo.woollacott@wits.ac.za

This paper presents an example of engineering educational research as a means of improving the effectiveness of teaching through the identification of student challenges and the subsequent creation of educational strategies to mediate gaps and enhance student mastery of an engineering topic. Its specific focus is student assessment error analysis, an under-researched field in engineering education. The context of the study was a cohort of 133 students taking a reactor design module in a chemical engineering degree programme. The study demonstrates how insights gained from a stringent data analysis of student test errors combined with teaching reflections can inform pedagogical redesign of a course. It discusses the utility of the research methodology used and how it can be deployed as a strategy in research-led engineering education.

Keywords: engineering education; student-centred education; evidence-based education; reactor design; student errors; student assessment error analysis

1. Introduction

This paper is part of a series that presents case studies in research-led education by which we mean "using educational research to inform the design of one's courses" [1, p. 89]. This approach to teaching has gained considerable traction in higher education to the extent that it is becoming increasingly recognized as an effective general strategy for improving the quality and effectiveness of teaching [1–5]. The cases presented in this series arose from the deployment of this strategy in a school of engineering by involving lecturers in the school in collaborative educational research with a colleague on the staff experienced in educational research. Each paper in the series focuses on a particular kind of investigation. The focus of this paper is the analysis of errors that students make when writing tests or examinations. In this context, we define an error very broadly as a deviation from the relevant norms, i.e. from what is considered within the relevant disciplinary community as being a valid or accepted conception, procedure, practice or argument.

The intention behind having students write a test is that their written work will reflect or give an indication, in the relevant context, of the status of their conceptual understanding and their ability to perform at the expected level of competence. It is our postulate that in the context of writing tests, the errors students make provide a window into not only what is being assessed but also into their thinking, processing, learning, and how their learning and development has responded to teaching. Accordingly, we argue that an analysis of such errors is potentially a useful way for a teacher to gain insights that can inform or help refine the pedagogies they are using or would like to use. This is demonstrated in the study presented in this paper as well as by the brief review of literature presented shortly.

2. The study

2.1 Context of the study

The context of the study was a third year chemical engineering design principles course offered by a South African University. The module on reaction engineering constituted 25% of the course. In brief, the module was taught in a conventional manner—two lecture periods per week over a single semester with a weekly afternoon tutorial. The numbers of students registered for the course in 2014 was 133.

In previous years, students had found the topic to be difficult and the performance of the 2014 cohort as a whole was particularly poor—the failure rate for the module was 56%. It was not clear to the course coordinator which aspects of the module caused students greatest difficulty and what factors contributed to their generally poor performance in the mid-term and mid-year tests. The study aimed to gain a clearer understanding of what these difficulties and factors were so that appropriate pedagogical measures could be taken. As a first step in that direction, the errors which the students had made in a written test were investigated in detail. To guide the investigation the following review of literature was undertaken.

2.2 Review of literature

The idea of learning from students' errors has a long history. It has been implemented, for example, in

* Accepted 30 December 2016.

engineering education [e.g. 6–8], in computer science education [e.g. 9, 10], elearning [10], and, most prominently, in mathematics education [e.g. 11–13] and in science education (see, for example, the 8400 references in [14]). Research of this kind falls broadly into two camps: students learning from the errors they make (e.g. [9, 15]), or teachers learning from student errors (e.g. [11, 13]). In the latter case, which is the area relevant to this paper, the most extensive body of work that has been carried out is in the fields of mathematics and science education and in regard to students' conceptual deviations variously labelled 'misconceptions', or 'alternative' or 'naïve' or 'non-scientific' conceptions [6]. It has been shown that many "misconceptions and the errors they produce are remarkably persistent and similar across contexts, independent of curricula or teaching methods, and thus can be seen as normal and possibly necessary steps in the development of mature concepts" [11, p. 221]. The pedagogical implications of this observation are obvious; knowledge about the nature and causes of such misconceptions is not just useful but vital for teaching that aims to facilitate in students a solid and integrated grasp of the body of concepts that undergird a discipline.

There has been a very large body of research, particularly in science education, that has developed inventories ('concept inventories') of the range of misconceptions or alternative conceptions associated with particular concepts [e.g. 10, 16, 17] and has researched the use of these inventories for diagnostic and assessment purposes [e.g. 10, 13, 16]. This kind of research attempts to identify the conceptual variations associated with any given concept and then to ask questions such as "what could the learner be thinking in order to make the error; how can we see the error from the learner's perspective; and how might the error make sense to the learner, even if not to the teachers" [11, p. 229].

Insights that are forthcoming from this kind of 'misconceptions research' can enrich a teacher's knowledge of how best to 'teach' the concepts in question [e.g. 7, 11, 12], i.e. it enriches their pedagogical content knowledge [18, 19]. This in turn can have multiple benefits for student learning such as providing insights that help teachers to better understand the roots of persistent misunderstandings among their students [e.g. 7, 11, 13]; to engage more effectively with students in classroom discussions on a topic [e.g. 15]; to diagnose the difficulties their students are experiencing in their conceptual mastery of a topic [e.g. 13, 20]; and to modify pedagogies and design appropriate interventions [e.g. 7, 12, 14, 20].

Misconceptions are not the only factors behind the errors students make. Other factors that may be significant include, for example, difficulties in understanding and solving problems, and inadequacies in students' application skills. In principle, errors that derive from these other factors can be researched in a manner similar to research on misconceptions and can have a similar impact on teaching and learning. One significant difference from misconceptions research, however, is that the focus of the research must be broader. Misconceptions research selects a particular concept and investigates the variations in students' conceptions of that concept and the consequential pedagogical implications. In contrast, research on students' errors in general must start with the identification of the different types of errors they make and then proceed to analyse the nature and causes of those errors. Misconceptions research in engineering education is under-developed [6] and we were unable to find any literature on misconceptions relevant to reaction engineering that also had a broad focus on the full range of errors students might make in this area.

2.3 Research method

Our literature review did not reveal any particular research framework for investigating the full range of errors students make in tests. Accordingly, a grounded approach [21] was adopted for the study. The questions addressed were as follows.

(a) What kinds of errors did students make when answering test questions in this topic?
(b) What was the prevalence of these errors among the students' test scripts and what was the relationship between them and the marks achieved in the test?
(c) In what ways did these errors cluster and what do these clusters signify about the difficulties students had in mastering this topic?
(d) What are the pedagogical implications of the findings from (a) to (c)?

To address these questions, the investigation focused on the 2014 midterm test. The lecturer who had set and graded the test reviewed each test script a second time in some detail to identify the errors that the students had made, to code each kind of error using an appropriate descriptive phrase, and from this to develop an emergent list of error types. Once all the scripts had been reviewed, the list of error types was consolidated by collapsing similar error types into one. The error types each student had made were noted and the information was compiled into a matrix that mapped the type of errors made by each student in the test. The matrix was then augmented by adding the mark each student had achieved so that the relationship between error type and student development as reflected by performance in the test could be investigated. Univariate and multivariate General Linear Models were used to investigate this relationship.

The data was analyzed further to identify groupings (clusters) of error types that tended to occur together as it was envisaged that it might be better for interventions to address clusters of error types rather than individual error types. A hierarchical cluster analysis was performed on the data based on a Jaccard dissimilarity matrix [22]. Single linkage and complete linkage clustering algorithms were used. All statistical analyses were conducted in SAS [23].

3. Research findings

3.1 Error types

Twelve error types were identified and are listed in Table 1 by error number along with a descriptive code, an explanation and examples.

Figure 1a shows the distribution of the number of different types of error made by each student and Figure 1b the distribution of marks achieved by the students. As can be seen from Fig. 1a, 62% of the students made 3 or 4 of the error types identified, 24% made 5 or 6, while only 11% made 2 or less.

The prevalence of the occurrence of each of the

Table 1. Types of errors made by students in the 2014 mid-year test

	Descriptive Code	Explanation and examples
E1	Skipping calculation steps	Errors arising from skipping calculation steps (not writing down all the steps).
E2	Time constraints (not finishing)	Student did not complete the problems apparently because they ran out of time.
E3	Not understanding/not reading the question	Student did not understand the question or did not read the given information accurately.
E4	Poor grasp of definitions/ theory	Error arising apparently from not knowing the relevant definition or theory with regard to, for example, elementary kinetics, order of reaction, equilibrium constant, void factor, density, and limiting reagents.
E5	Did not remember formulas	Self explanatory.
E6	Transferring and calculation errors	Calculation mistakes; errors arising from incorrect copying of numbers or equations from previous lines.
E7	Does not understand conversion	Error arising from a poor grasp or application of the concept of conversion (for example, not defining it appropriately; how conversion works using the definitions provided; applying the concept when reactors are in series)
E8	Does not understand stoichiometric table concepts	Error arising from a poor grasp of stoichiometric tables (for example, how to setup the stoichiometric table; applying it for reactors and not for streams).
E9	Gave up when faced with difficulties	Student 'gave up' when faced with intermediate errors or complicated maths.
E10	Integration error/s	Calculus error with respect to integration (for example struggling with the idea of changing variables over the area of integration in design equations).
E11	Fundamental errors	Fundamental conceptual errors (such as applying the ideal gas law to liquids).
E12	Errors with units	For example, not paying attention to the units of the values given, and using different units in the same equation.

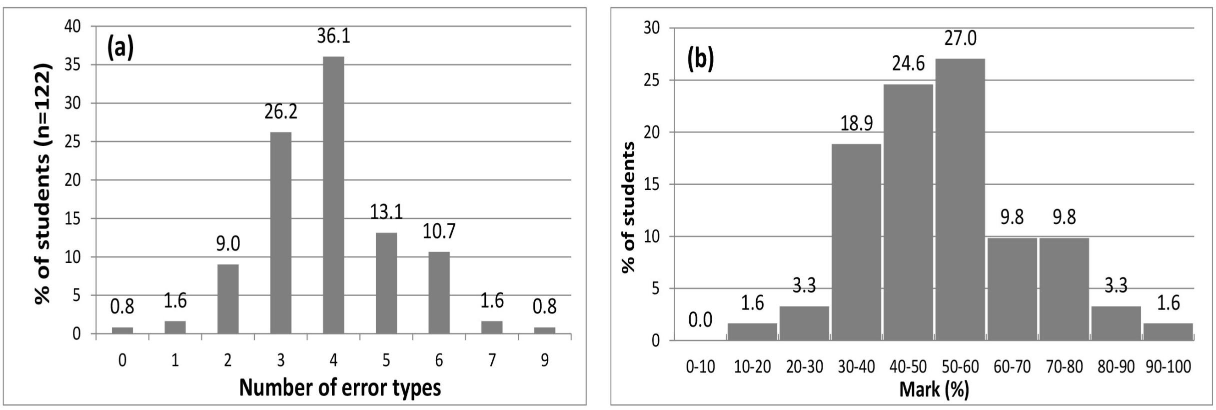

Fig. 1. Descriptive statistics (n = 122). (a) Distribution of the number of different error types made by a student. (b) Distribution of marks achieved by students.

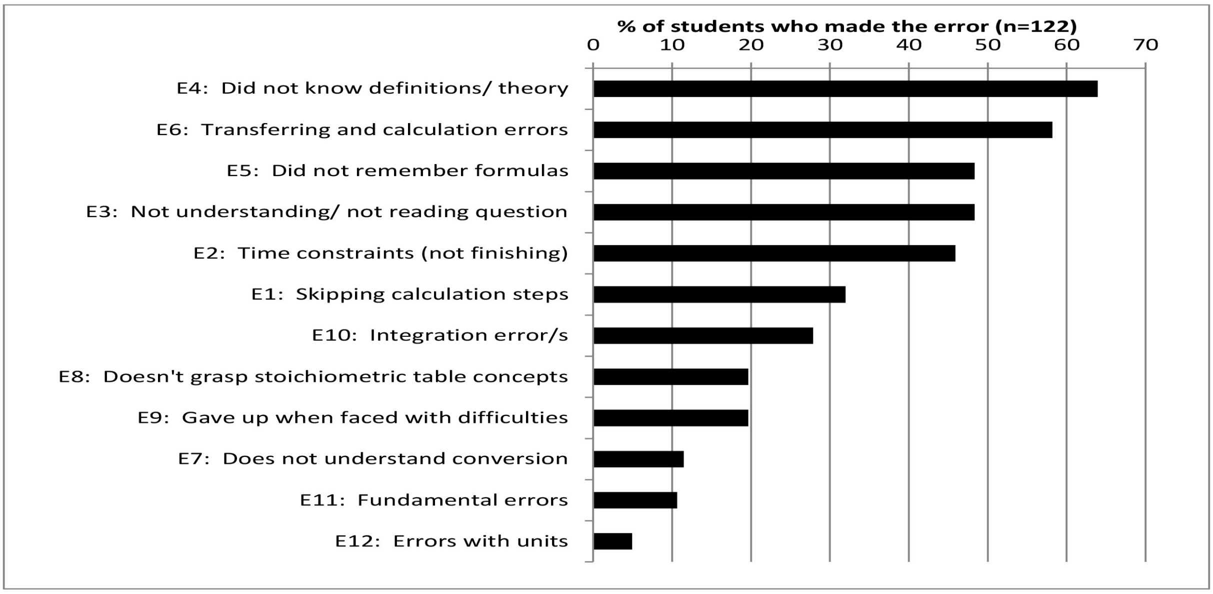

Fig. 2. Prevalence of the occurrence of the different error types (n = 122).

different error types is shown in Fig. 2. The most common errors were E4, poor grasp of definitions/ theory (associated with 64% of the students) and E6, transferring and calculation errors (58%). The least common error was E12, the incorrect use of or attention to units (5%).

3.2 Impact of error types on test marks

The distribution of the marks students achieved in the test is shown in Fig. 1b. The marks ranged from 12 to 95% with a mean of 51.2% and a standard deviation of 15.5%. (Note that the pass mark was 50% and that 75% or more was a distinction.) The size of the sample and the structure of the data (there was no strong pairwise association between any of the error types) allowed the application of a multivariate General Linear Model, Equation (1), to investigate the relationship between the error types and the test mark.

$$\text{Test Mark} = \beta_0 + \beta_1(\text{error type E1}) + \beta_2(\text{error type E2}) + \ldots + \beta_{12}(\text{error type E12}) + \varepsilon \quad (1)$$

where β_0 , β_1, . . . β_{12} are coefficients in the model, and ε is the residual error.

The results of this analysis are shown in Table 2. From a statistical perspective, the first 7 error types in the table were found to be significantly related to the test mark. Also shown in the table is an estimate of the impact of making a particular type of error as suggested by the General Linear Model, Equation (1). This estimate is expressed as the increase in the test mark that would be achieved by not making the

Table 2. The multivariate relationship between error type and test mark

	Error Type	F Value	p-value	Estimated increase in Test Mark if error type not present: Estimate	95% CI	Prevalence of error (%)
Statistically significant relationships						
E12	Errors with units	6.53	0.012	14.0	3.3 to 24.7	4.9
E4	Poor grasp of definitions/theory	30.60	<0.0001	13.0	8.4 to 17.6	63.9
E8	Doesn't grasp stoichiometric table concepts	9.12	0.0031	8.4	2.9 to 13.8	19.7
E5	Did not know/remember formulas	13.16	0.0004	8.3	3.8 to 12.7	48.4
E3	Not understanding/not reading question	13.63	0.0003	8.2	3.9 to 12.6	48.4
E2	Time constraints (not finishing)	11.92	0.0008	8.1	3.5 to 12.7	45.9
E1	Skipping calculation steps	5.02	0.027	5.3	0.7 to 9.9	32.0
Statistically non-significant relationships						
E10	Integration error/s	3.44	0.066	4.7	–0.3 to 9.7	27.9
E9	Gave up when faced with difficulties	2.30	0.13	4.2	–1.2 to 9.7	19.7
E11	Fundamental errors	0.04	0.85	0.7	–6.6 to 8.0	10.7
E7	Does not understand conversion	0.01	0.94	0.3	–6.7 to 7.2	11.5
E6	Transferring and calculation errors	0.30	0.58	–1.3	–5.9 to 3.3	58.2

error and is derived from the difference between the test mark achieved by those who did not make that type of error and the mark achieved by those who did, controlling for the effect of the other error types. The error-types are listed in decreasing order of this estimate. To assess the practical significance of these estimates consideration must be given to the 95% confidence interval (CI) of the estimate as well as to the prevalence of the error among the students; both are included in Table 2.

To facilitate the assessment of the practical significance of the information in Table 2, Fig. 3 plots the estimate of the impact on the test mark of not making a particular error type against the prevalence of that error type. In the first place, the figure highlights two error types as 'outliers'—E12, errors with units, and E6, transferring and calculation errors. The impact of the first of these error types on test marks appears to be considerable (the estimated impact is the largest, 14%) but is variable (the confidence interval has a very large range) with a low prevalence among the students (the lowest by some margin). The low prevalence suggests that the error type may be more of a slip or a mistake than a conceptual problem. If this conclusion is valid, little pedagogical attention is required to help students to avoid this type of error beyond making them aware that it can have a significant impact on the accuracy of calculations and perhaps emphasizing the need to be meticulous in their attention to the use of correct units in those calculations.

The second error type highlighted as an 'outlier' in Fig. 3 (i.e. E6, transferring and calculation errors) had a high prevalence among the students but very little impact on test marks. When the confidence interval is taken into consideration, –5.9 to 3.3%, it is apparent that the impact on marks is not significantly different to zero. On the one hand, this suggests that the grading emphasized method over calculation accuracy, which is an assessment decision that may or may not be appropriate in the context. On the other hand, the high prevalence of the error type is a cause for concern calling for some form of pedagogical intervention as discussed shortly.

The stand out error type in both the figure and the table is error-type E4—poor grasp of definitions/ theory. It has the second highest estimated impact on test marks, 13%, and the highest prevalence among students (64% of students made this kind of error). The pedagogical implications of this error-type are ambiguous. On the one hand, it is apparent from the examples of this error type that the errors have to do with concepts that are relatively basic, some taught or reinforced in the course, but most taught in earlier years of the degree programme (and in some cases taught at school!). The implication is that these basic concepts or definitions have been inadequately grasped by many of the students. However, it is also possible that the errors were simply mistakes or slips resulting from time or test pressure. Another possibility is that the grasp of the concepts/definitions concerned was not sufficiently robust for accurate application in a wide range of contexts and, more specifically, in the context of the problems being solved in the test. An additional possibility is that the nature or language of the problems set in the tests created uncertainties in the minds of some students leading to a misapplication of the concepts or definitions. These issues will be taken up again in the discussion of the clustering of the error types.

Before that discussion, however, it is appropriate to consider a second, possibly related, error type, E5 'did not know/remember formulas'. This is listed fourth in Table 2 with a prevalence of 48% and is associated with an estimated 8.3% reduction in the test mark when the error is made. Two pedagogical issues appear to require attention here. The first is that further investigation or analysis is called for to ascertain which aspects of the theory (and the associated definitions and formulae) are proving

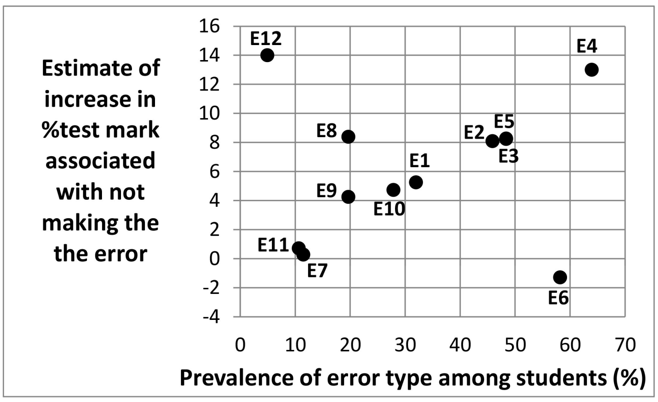

Fig. 3. Prevalence of an error type compared to its impact on test marks.

problematic for students to master, and thereafter to reflect on how those aspects are taught and reinforced. The second issue is the attention given in the course (and in assessments) to the remembering of formulae. For example, the table suggests that the marks of many students might increase by about 8% if students knew they were not required to remember formulae and that a list of relevant formulae would always be provided in the test question paper. Whether or not this would conflict with other educational or assessment strategies, it does suggest that this aspect of the course requires further reflection and perhaps investigation and modification.

To inform further deliberations on the pedagogical implications of the information in Table 2 and Fig. 3, it is useful to consider how the error types clustered. This is addressed next.

3.3 Clustering of the error-types

A cluster analysis was conducted to investigate the extent to which error types tended to occur together in students' scripts. The results of the hierarchical clustering are presented in the form of a dendrogram [22]. Fig. 4 shows the dendrogram from the complete linkage clustering algorithm (the results from the single linkage clustering algorithm were very similar). The 12 error types are shown on the left. As one moves to the right across the dendrogram, the progressive clustering of the error types is shown until, on the far right, only one cluster (consisting of all 12 error types) remains. Interpretation of the cluster structure is typically made where there is no change in the cluster structure for a considerable distance along the dendrogram. In this case, four distinct clusters (groupings) of error types are indicated. These are labelled and discussed after the diagram.

Cluster 1: Generic execution errors—E1 to E6. The error types in this cluster are, firstly, skipping calculation steps (E1); running out of time (E2); not understanding the question or not reading it accurately (E3); and making transferring and calculation errors (E6). These have to do with the process of understanding and solving 'calculation-heavy' problems under time pressure and are generic in nature in that they are not specific to the topic or in fact to any particular disciplinary topic.

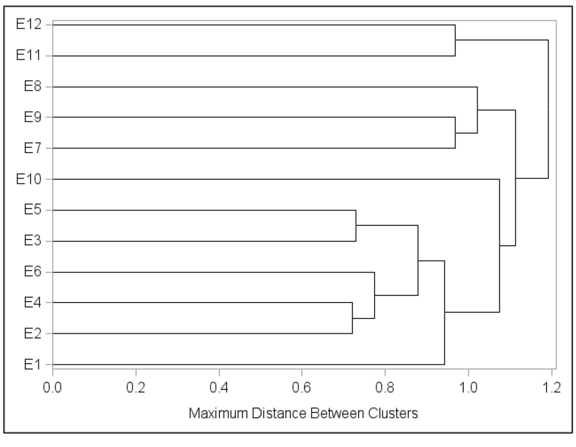

Fig. 4. Complete linkage cluster analysis of the error types E1 to E12 showing which error types tended to occur together.

The cluster also includes 'a poor grasp of basic definitions or theory' (E4); and 'not remembering the basic formulae' (E5). A number of possible ways of interpreting the pedagogical significance of these two error types have already been considered. Some of those possibilities related to 'slips', or misapplication of the concepts/definitions in the context of the solving of the test problem. The fact that these two error types cluster with the four 'execution errors' just described lends support to such an interpretation (although it does not rule out the other possibilities). Alternatively, it could be argued that the clustering of these two error types with error-types that are clearly examples of generic execution errors makes sense.

Figure 2 shows that the six error types in this cluster are the six most commonly made errors and that the prevalence of these error types is very high—between a third and two thirds of the students made these types of error. These observations have a number of pedagogical implications. The most obvious is that accuracy in executing calculations is a major problem among these students and that many of them may be failing the course because of this 'execution' problem rather than because they lack mastery of the subject. To counter this problem, some form of intervention or pedagogical modification is indicated and further research into why these execution problems occur should be done to guide the design of the modifications. An obvious modification is to increase the amount of attention and course time given to the solving of problems. Further, the cluster and its prevalence raises questions about why the earlier years in the degree programme have failed to develop in the students a higher level of test-taking and problem-solving competency. In addition, it is evident that consideration should be given to the form of the assessment used, particularly with regard to the need to memorize formulae, and to the extent to which some of the execution difficulties noted may be being induced or exacerbated by that form of assessment.

Cluster 2: Generic integration errors—E10. This cluster consisted of only one error type (E10) which has to do with the application of calculus. It is a generic error type in that it is not topic specific. However, unlike the generic execution error types of Cluster 1, it is rooted in a specific skill inadequacy with regard to the practical application of mathematics in an engineering context. According to the dendrogram this error type may be considered to stand alone or to be associated with Cluster 1.

The prevalence of this error-type is high—nearly a third of the students made errors of this kind. The pedagogical implications of this observation point in two directions; to problems in the prior development of students' applied skills in this area of mathematics; and to the need to give extra attention in the course to rehearsing or reinforcing this skill as it is applied in reactor design.

Cluster 3: Conceptual shortcomings—E7 to 9. The two error types in this cluster—'an inadequate understanding of the concept of conversion' (E7), and 'does not understand stoichiometric table concepts' (E8)—are topic specific and point to an aspect of the course where specific interventions and further research are indicated. This cluster also includes error type E9—'giving up when faced with difficulties' particularly with regard to complex maths. Intuitively the association of this error type with the two conceptually based error types in this cluster makes sense in that the application of the two concepts leads to fairly complex mathematical expressions in reactor design and these could become overwhelmingly complex or difficult if the relevant concepts are misapplied and, as a result, could lead to difficulties which may make students 'give up' some way into a problem in a test.

From a pedagogical perspective, error type E8—inadequate grasp of stoichiometric table concepts – appears to require more critical attention than the other conceptually based error type, inadequate grasp of conversion (E7). It has a greater prevalence (19.7% compared to 11.5%) and a greater impact on test marks (8.4% compared to 0.3%). The difference in impact may be due to a larger number of ways in which errors can be made with regard to the stoichiometric table as compared to the application of conversion principles or simply that the former is more difficult for the students to grasp and apply accurately than the latter. From the literature review it is evident that a further investigation into student misconceptions in these two areas may prove fruitful.

Cluster 4: Fundamental conceptual errors—E11 to 12. The two error types in this cluster—fundamental errors (E11) and errors with units (E12)—are fundamental in nature in that they relate to foundational concepts or skills not taught in the course but assumed to be have already been mastered by the students. These were among the least common error types identified in the student scripts. As such, little pedagogical attention seems necessary beyond bringing to the attention of the students at appropriate points in the course the nature of these errors and the kind of impact they can have on problem solving in the topic.

4. Discussion

Up to this point in the paper, a panorama of findings and insights has been presented. It is now appropriate to discuss the broader significance of these and to do so at three levels, namely (1) pedagogy—the impact of the study on the teaching of the reactor design module; (2) research methodology—what the study has to say about the methodology used; and (3) research-led engineering education—what the study has to say about how educational research can inform engineering education and impact the quality of teaching and learning.

4.1 The impact of the study on the teaching of the reactor design module

The insights obtained by identifying and analysing the errors the students made in the mid-year test were both specific and strategic in nature. In the first place, the twelve error types that were identified had specific pedagogical implications in that they each suggested one or more shortcomings in the student's understanding or skill along with possible modifications to teaching that could be implemented to address those shortcomings. This 'interpretation' of the error types was not always unambiguous, however, and sometimes needed to be clarified by deeper analysis as discussed in the next section. These specific pedagogical implications were described earlier and are not repeated here. Together they enriched the 'pedagogical content knowledge' [18, 19] of the lecturer in that her knowledge of the students' needs and responses to her teaching was enhanced. In addition, her reflections on what she had learned generated new ideas about how to teach aspects of the subject and also led to the strategic shifts described next.

At a strategic level, it was apparent that two shifts should be made in the way the topic was taught. The first was in the area of knowledge application and problem solving skill. The study findings highlighted 'generic execution errors' in the solving of problems as a major shortcoming in the competency of many students as reflected in their performance in the test. There was also a credible indication that the apparent shortcomings in the conceptual grasp of many students may have been more related to the application of that knowledge than to a lack of conceptual understanding. Accordingly, this suggests a shift is needed in the way the course is designed to facilitate students' development of a more robust skill in the solving of reactor design problems.

The second area in which a strategic shift in pedagogy is indicated is in the area of theoretical knowledge. Although shortcomings in knowledge application appeared to be a major problem, as already mentioned, there were also strong indications that attention should be given to how the course facilitated the students' conceptual development with regard to key concepts in reactor design particularly those related to the stoichiometric table and its conceptual application.

In the short term, these shifts in strategy were implemented by modifying the focus of teaching in the lecture periods and the structure of the weekly afternoon tutorial sessions. One of the weekly lecture periods became devoted to theory and 'application exercises' and the other to solving 'application problems'. The distinction between 'exercises' and 'problems' is that the former involves straight forward application of theoretical principles while the latter consist of more complex and realistic examples of reactor design situations. In this way, the theory is introduced and reinforced in the context of problem solving and the solving of problems is addressed in the context of mastering key concepts and is scaffolded from easier to more demanding and realistic situations. The weekly tutorial sessions were also restructured so that they were tightly integrated with the content of the lecture periods and provided deeper reinforcement of the students' developing application skills in reactor design. The implementation of this strategy is ongoing and it is currently too early to comment on its impact.

In the longer term, consideration is being given to the development or use of on-line 'lecturettes' and videoed solutions to problems, and to a deeper integration of theory and problem solving through, for example, the use of a problem-based learning approach.

4.2 The interpretation of student errors as a research methodology in engineering education

Assessments are designed to evaluate the degree to which each student individually has mastered a particular topic or developed a particular skill. The research methodology described in this paper uses the written work handed in for assessment purposes with a different objective in mind, namely to assess the appropriateness and effectiveness of 'teaching' (i.e. of the educational strategy used to facilitate student learning). The written work submitted by a class for grading is treated as a collective reflection of how the students have responded to that 'teaching'. It is analysed to identify and classify the types of errors made, the prevalence of those errors; their impact on test marks; and how they cluster together in the test scripts. The findings of this analysis are then 'interpreted' for their pedagogical significance. By this we mean that possible reasons for the different errors are considered along with the associated pedagogical implications if any. The methodology has been described and illustrated in some detail in the paper. However, more needs to be said about the interpretation of the errors because, up to this point, such interpretation has been largely heuristic in nature.

The first point that emerges from a consideration of the interpretations described in the study is that there may be a number of reasons why students made the errors they did and therefore a number of indications of possible pedagogical measures that may be implemented to enhance student mastery and so reduce the occurrence of those errors. This ambiguity in the pedagogical implications associated with each type of error can be reduced and perhaps eliminated when the prevalence, impact and clustering associated with an error type is taken into consideration. The latter appears to be particularly useful in this regard because of the way that it can draw into consideration the pedagogical implications associated with more than one type of error.

A second point that emerges from the study is that the analysis of errors can only give an indication of the pedagogical measures that might be appropriate. Anything more than this is essentially precluded by the interpretive nature of the analysis if no other data such as questionnaires or interviews is gathered from the students. For the same reason, it seems pointless to press the interpretation of the errors too far or too deeply. Despite this limitation, however, the indications that can be developed from the analysis can be very useful in informing the pedagogy employed or the redesign of the education strategies used in a course as the case presented in this paper has illustrated.

Finally, it is perhaps useful to classify the types of pedagogical indications that can emerge from an interpretation of students' errors. Reflection on the range of error types that were identified in the study suggests the following classification.

(a) *Mistakes*: These are errors that arise from performance 'slips' rather than from shortcomings of understanding or skill. A low prevalence of a particular type of error suggests that it may simply have been a mistake and therefore that no particular pedagogical intervention is indicated.

(b) *Application related errors*: These point to shortcomings in some aspect of problem solving skill. They may relate to the ability to interpret the circumstances of the problem appropriately; to choosing or developing an appropriate problem solving procedure; to accuracy in executing that procedure; to recognizing which concepts should be used to solve the problem; to the accurate application of that concept; or to some mix of all of these. The pedagogical implications of such errors are to pay more attention in the course to developing or reinforcing the appropriate skill in the students.

(c) *Concept related errors*: These point to shortcomings in the students' understanding of the concepts addressed in the course. Here the pedagogical indications are to modify how those particular concepts are 'taught' in the course and perhaps to research student misconceptions related to those concepts.

(d) *Curriculum related errors*: These point to shortcomings in the skills or understanding of the students that are not specific learning outcomes of a course and are presumed to have been mastered by the students prior to the course. As such, they derive from the general structure and effectiveness of the curriculum of the educational programme as a whole. Such errors have both short and longer-term implications. In the short term, modifications within the course may be indicated to compensate for the shortcomings in the curriculum. In the longer term, modifications to the wider curriculum are indicated.

In addition to the above classification of error types, it is evident from the case study presented in this paper that a distinction should be made between specific and strategic pedagogical indications.

(e) *Specific pedagogical indications*: These are indications that derive from specific errors and point to specific pedagogical measures to address the shortcomings associated with that type of error.

(f) *Strategic pedagogical indications*: These are indications that derive from a consideration of error types considered together as a whole and point to more wide-ranging shifts needed in the pedagogies and educational strategies employed in the course.

4.3 The study as an example of research-led engineering education

The ultimate objective of educational research is to improve the effectiveness of teaching and educational strategies as reflected in the quality of student learning by formulating pertinent questions and gathering and analysing credible evidence to answer those questions. The case study presented in this paper is an example of research of this kind. The questions posed arose from a teacher's concern to gain a better understanding of the factors behind the relatively poor academic performance of students taking a module in an engineering course. The research methodology implemented to answer those questions provided a range of insights that led to a number of modifications to the way the course was taught. Importantly, these insights were researched-based in contrast to being derived from anecdotal observations, reflections, or the teacher's intuitions. How the modifications impacted the effectiveness of teaching and learning is the subject of a further research project currently being undertaken.

To conclude the discussion of the case study as an example of research-led education, three points are worth making. The first is that the educational research conducted in the study generated new knowledge of two kinds: pedagogical content knowledge useful for the teacher who conducted the research and for other teachers in similar circumstances; and methodological knowledge. The methodology developed in this case study appears to be quite new in the field of analysing student errors for pedagogical purposes. It has a broader focus than misconceptions research which focuses on a selected concept and how students' conceptions deviate from the relevant norms. In contrast, the methodology developed in this study attempts to identify and analyse all the different kinds of errors students make in a test, and the factors and interrelations behind these, and includes all this information in the analysis. An innovative aspect of this methodology is the way that insights into the causes of student errors are sharpened by triangulating three sources of information—the prevalence of the errors among the students, the clustering of those errors, and their impact on test marks.

The second point to note is that the knowledge generated in the study has emerged from the depths of teaching practice, i.e. a teacher researching teaching and learning in the context of their own teaching practice. Woollacott [24] has pointed out that some kinds of educational knowledge can only be developed in this way.

The third point to make is that the full implementation of the methodology developed in this study takes time and the findings and reflections on those findings are likely to be available only in the semester or academic year after the research has been carried out. It may be possible to accelerate the analysis or conduct a partial analysis so that findings become available for implementation more timeously. However, this would require further research and development.

5. Conclusion

The case study presented in this paper has demonstrated, in a specific context, how the errors students make in summative tests provide a window into how their learning and development has responded to teaching. It has shown further how an analysis of these errors can be conducted and how the findings from such an analysis can inform or help teachers to refine the pedagogies they are using or would like to use. The approach that has been developed is novel and is broader than, and complements, other approaches that analyse student errors such as those used in research on student misconceptions. It is particularly useful for situations where a teacher is unsure about the exact nature of the difficulties students are experiencing in mastering a topic. The methodology developed in this paper draws from three sources of data gathered from student tests scripts: the prevalence of the different types of errors the students made, how these errors cluster, and the grades the students achieved. It provides a way of classifying the different kinds of errors the students make in a topic, the possible reasons behind those errors, and thereby points to pedagogical measures that might be taken to improve the effectiveness of teaching and the quality of student learning in that topic.

References

1. K. Valter and G. Akerlind, Introducing students to ways of thinking and acting like a researcher: A case study of research-led education in the sciences, *International Journal of Teaching and Learning in Higher Education*, **22**, 2010, pp. 89–97
2. R. Osman and H. Venkat, *Research-led Teacher Education*, Pearson, Cape Town, 2012.
3. B. Zamorski, Research-led teaching and learning in higher education: A case, *Teaching in Higher Education*, **7**, 2002, pp. 411–427.
4. N. Entwistle and J. Mighty, Taking stock: An overview of research findings, *Research on Teaching and Learning in Higher Education* [Online], 2010, [Accessed Feb 2016].
5. K. Trigwell, Research-based university teaching: A response to Noel Entwistle, *Psychology of Education Review*, **26**, 2002, pp. 23–24.
6. T. V. Goris and M. J. Dyrenfurth, Concepts and misconceptions in engineering, technology and science. Overview of research literature, *2012 IL/IN Sectional Conference*, Valparaiso University, Valparaiso, Indiana, American Society for Engineering Education, 2016, pp. 1–16.
7. S. Krause, J. Kelly and D. Baker, Addressing Misconceptions and Knowledge Gaps in Restructuring of Atomic Bonding Content in a Materials Course to Enhance Student Conceptual Change, *ASAA/IEEE Frontiers in Education Conference*, 2010 Washington, DC.
8. R. A. Streveler, S. Brown, G. L. Herman and D. Montfort, Conceptual change and misconceptions in engineering education. In *Cambridge Handbook of Engineering Education Research*, Cambridge, U.K., Cambridge University Press, 2015.
9. D. Ginat and R. Shmallo, Constructive use of errors in teaching CS1. *SIGCSE'13: The 44th ACM Technical Symposium on Computer Science Education*, Denver, Colorado, USA, 2013, pp. 353-358.
10. J. Priem, 2010. Fail better: toward a taxonomy of e-learning error, *Journal of educational computing research*, **43**, 2010, pp. 377–397.
11. K. Brodie, Learning about learner errors in professional learning communities, *Educ. Stud. Math.*, **85**, 2014, pp. 221–239.
12. A. Peng and Z. Luo, A framework for examining mathematics teacher knowledge as used in error analysis, *For the Learning of Mathematics*, **29**, 2009, pp. 22–25.
13. D. Treagust, Diagnostic assessment in science as a means to improving teaching, learning and retention, *The Assessment in Science Teaching and Learning Symposium*, Sydney, 2006, pp 1–9.
14. R. Duit, *Students' alternative frameworks and science education* [Online]. Available: http://archiv.ipn.uni-kiel.de/stcse/, 2009. [Accessed 9 December 2016].
15. J. Lannin, B. Townsend and D. Barker, The reflective cycle of student error analysis, *For the Learning of Mathematics*, **26**, 2006, pp. 33–38.
16. C. D. Flynn, C. I. Davidson and S. Dotger, From assessment to intervention: Conceptual understanding of rate and accumulation processes, *ASEE Annual Conference and Exposition*, Paper 16989, 2016.
17. S. Nozaki, N. E. Study, H. M. Steinhauer, S. A. Sorby, M. A. Sadowski and R. L. Miller, Development of a concept inventory for engineering design graphics, *Frontiers in Education Conference (FIE)*, IEEE, 2016.
18. A. Berry, J. Loughran and J. H. Van Driel, Revisiting the roots of pedagogical content knowledge, *International Journal of Science Education*, **30**, 2008, pp. 1271–1279.
19. L. S. Shulman, Those who understand: Knowledge growth in teaching, *Educational Researcher*, **15,** 1986, pp. 4–14.
20. W. H. Cliff, Case study analysis and the remediation of misconceptions about respiratory physiology, *Advances in Physiology Education*, **30**, 2006, pp. 215–223.
21. L. Cohen, L. Manion and K. Morrison, *Research Methods in Education*, London, RoutledgeFalmer, 2007.
22. B. S. Everitt, S. Landau, M. Leese and D. Stahl, *Cluster Analysis*, Wiley, New York, 2011.
23. SAS, *SAS Software, version 9.3 for Windows*, SAS Institute Inc. 2002–2010.
24. L. Woollacott, A case study in the development of an innovatve pedagogy. *In:* S. Booth and L Woollacott (eds.) *The scholarship of teaching and learning in higher education: On its constitution and transformative potential*, Stellenbosch, South Africa: Sun Media Metro, 2015, pp. 55–92.

Laurie Woollacott is currently a visiting associate professor at the University of the Witwatersrand, Johannesburg in the School of Chemical and Metallurgical Engineering. From a background of industrial experience in minerals processing and metallurgical design he joined the school in 1983. He has an MSc in Engineering and a PhD in higher education. He has co-authored a textbook on extractive metallurgy, co-edited a book on the Scholarship of Teaching and Learning, and has numerous publications in both engineering and educational research.

Lizelle van Dyk is a senior lecturer at the University of the Witwatersrand, Johannesburg in the School of Chemical and Metallurgical Engineering. She has an MSc in Extractive Metallurgy and a PhD in Chemical Engineering from the University of Stellenbosch, South Africa. She has published in the field of extractive metallurgy and membrane science.