A Design Science Approach to Assessment of Multi-Steps Questions in Mathematics

Hussein Genemo

College of Engineering & Science

Victoria University Melbourne

Australia

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

August 2019

Abstract

Academic assessment that is used to assist in producing knowledge having a lasting effect should be the concern of all stakeholders. Subjective assessment contributes to achieving this type of knowledge. In particular for multi-steps questions (MSQ) human assessors play important roles. This evaluation includes assessing solution strategies, working produced from employing those strategies, and the final answer. This assessment can reveal levels of conceptual understanding displayed by students. Subjective assessment gives options to students to express their understanding of the topic. However, when computer-aided assessment (CAA) systems are used, there is much less subjectivity. These systems subdivide MSQ into stub-steps, and students can provide one step's answer without showing their workings. This technique is similar to assessing objective questions such as multiple-choice questions (MCQ) which do not, generally, examine conceptual understanding. The inability of CAA systems to assess solutions of MSQ is considered a complex problem. To solve a complex problem, accumulating knowledge about the problem, producing solutions and acquiring ability to apply those solutions are required. Design science research (DSR), which is the paradigm to investigate complex problems- in information systems research and design science - has been used in this research. Two hundred and fifty-eight student scripts containing working with solutions of MSQ were analysed using an inductive qualitative content analysis and the quantitative survey approach. The findings from the analyses of scripts were re-analysed iteratively to produce the knowledge that contributed to the understanding of the research problem and producing solutions. This knowledge includes types of solution strategies and student errors that can also be used in designing questions to be used in CAA systems for assessing objective questions. The methodologies that were used in producing this knowledge could also be used in similar disciplines to produce similar output. A significant contribution of this research is to analyse student workings for extracting solution strategies, and this helps understanding ways of solving MSQ and obtaining these strategies. This approach does not appear to have been used elsewhere. The questionnaire that was created for measuring the significance of types of strategies and student errors is an innovative instrument. It has not been used previously and can be adapted for similar studies. The information in the analyses of student workings also provides directions when solutions of MSQ are assessed by humans. Furthermore, all the above information as well as domain knowledge was used to develop DSR constructs and models artefacts. These models represent processes and data in the student workings, which was divided into sub-tasks with each task being represented by one model. The implementation of a model in a DSR prototype artefact accomplishes the assessment of that sub-task. One of these models was implemented, overcoming difficulty in designing user interfaces that can be used without disclosing the domain knowledge students are examined. These user interfaces show an innovative way of extracting conceptual understanding of topics. Furthermore, these models are highly generalisable to a very broad class of problems.

Doctor of Philosophy Declaration

"I, Hussein Genemo, declare that the PhD thesis entitled *A design science approach to assessment of multi-step questions in mathematics* is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

Signature [redacted] Date: August 8, 2019

Relevant Publications to this thesis

- Miah, SJ and Genemo, H. (2016). A Design Science Research Methodology for Expert Systems Development, Australasian Journal of Information Systems, 20, 1**-29**

- Genemo, H. and Miah, SJ. (2015). Assessment Issues in Mathematics: Design Science Approach for developing an Expert-System Based Solution, In the proceedings of the 26th Australasian Conference on Information Systems (ACIS 2015), Adelaide, Australia.
- Genemo, H., Miah, SJ. and McAndrew, A. (2016). A design science research methodology for developing a computer-aided assessment approach using method marking concept, Education and Information Technologies, 21(6), 1769–1784

Acknowledgement

In this thesis, I would like to acknowledge the supervision guidance and support by principal supervisor Associate Professor Shah Miah and associate supervisor Associate Professor Alasdair McAndrew.

Tables of contents

List of figures

List of tables

Chapter 1 Introduction

Assessment, according to Messerschmidt (2016), “is the process of systematic gathering, interpretation, and use of information about student cognitive, behavioural and attitudinal outcomes for purposes of improvement” (p. 2). However, in this thesis the focus is on academic assessment in relation to e-assessment systems for mathematics and assessing subjective questions. Authentic assessment is used to measure the comprehension of a concept or concepts so that actions can be taken, or feedback can be given based on the outcome of the assessment. Conceptual understanding of a subject is achieved generally by using relevant assessment of which subjective assessment is one. This conceptual understanding is also related to the type of approach to learning. The combinations of this learning approach and the proper assessment design results in generating the required conceptual understanding. Regarding this subject, Gijbels and Dochy (2006) stated:

> The literature on students’ approaches to learning suggests that ‘deep’ approaches to learning are encouraged more effectively by using assessment methods and teaching practices which aim at deep learning and conceptual understanding, rather than by trying to discourage a ‘surface’ approach to learning (Trigwell & Prosser, 1991). (pp. 399-400)

Generally, this type of assessment uses multi-steps questions (MSQ). In this type of questions, students are required to show their workings to display their conceptual understanding of the subject/topic. This type of questions can be traditionally assessed by humans. These assessors evaluate strategies in the working as well as answers using their own judgements guided by a rubric to award marks. However, even when traditionally assessed there can be problems in awarding marks (scoring). The inconsistency in the scoring of subjective working is present due to some examiners differences in what they think is important (White, 2019).

The difficulty of assessing student working in an MSQ becomes more obvious when e-assessment systems are used, as confirmed by the literature review. Therefore, further review of relevant literature and a study of an appropriate context is important to produce the knowledge that can be used in understanding these systems' inability of subjective assessment, as well as developing design science research (DSR) artefacts to be used in the studied context. Furthermore, doing further study on these DSR artefacts would help the generalisation of the knowledge that will be produced in this research.

The context that will be analysed is student workings in examinations scripts. These workings contain solutions of an MSQ. The workings will be analysed to extract solution strategies and student errors, and other related information to build the knowledge that has been mentioned above.

Gregor and Hevner (2013) suggest that the introduction section of a DSR project should include (1) an unambiguous statement showing the relevance of the research problem to real-world practice, (2) the research purpose and scope and (3) the contribution to be made "to practice and knowledge/theory" (p. 349). Hence, in this chapter, the research background and research problem issue will be discussed in section 1.1. Based on the outcome of the discussion, section 1.2 discusses this research purpose and scope. Using the arguments made in sections 1.1 and 1.2, section 1.3 overviews this study's 'research process'. The significance of the research is detailed in section 1.4 followed by section 1.5 that presents the thesis' structure with the brief summary of each chapter with the exception of this chapter. Section 1.6 presents the summary of this chapter.

1.1 Research background and research problem

The aims of academic assessment are to achieve certain goals which are “the outcomes that students are required to perform at the completion of the instruction” (Messerschmidt, 2016, p. 2). For these outcomes to be reflective of student understanding, Rose et al. (2018) assert that giving options to students to express what they understand and allowing them to communicate the way they understand derives “more accurate assessment results” (p. 168). Rose et al.’s argument indicates such types of assessment should be a subjective. Liu and Jansen (2015) also regard that the responses of subjective questions are based on views and skills of individuals. Subjective questions are similar to open questions which, for example, “encourage patients to fill in the gaps with respect to their feelings, thoughts, and situations” (McCarthy Veach, LeRoy, & Callanan, 2018, p. 12).

In some subjects such as mathematics, subjective assessment asks students to show their workings that involve using a strategy that might or might not be effective. According to Boekaerts, Seegers, and Vermeer (1995) “an effective strategy is connected with a high degree of confidence, or an ineffective strategy is connected with a high degree of doubt” (p. 257). When awarding marks to this type of working, what is important or measured is the degree of the effectiveness of the strategy that is used. This type of marking is based not only on awarding marks to the final answer but also includes awarding marking to the method (strategy) that is used.

Online or computer-aided assessment (CAA) systems are capable of assessing objective questions in comparison with the subjective questions (Hu & Xia, 2010). However objective questions are not adequate to examine sufficient knowledge. On the other hand, these systems show inability or difficulty to assess subjective questions (Hu & Xia, 2010). CAA systems use different approaches to award marks to solutions of subjective questions. Perfect (2015) reports some approaches that have been used by these systems. One method is to assess the final answer. The other method is dividing the questions into one step sub questions and give two options to students. Option one is to present one step at a time and expect students to produce each step's answer in turn. Option two differs only in that the solution of the first step is shown to students and they are expected to provide answers for the subsequent steps of the question. Beevers, Wild, McGuine, Fiddes, and Youngson (1999), Livne, Livne, and Wight (2007), and Sangwin, Cazes, Lee, and Wong (2009) reported the CAA systems using the first option. The other option of assessing student workings- using CAA systems- was reported by Ashton, Beevers, Korabinski, and Youngson (2006), Beevers et al. (1999) and Lawson (2012).

Hu and Xia (2010) have reported about the ability of CAA systems to assess objective questions but face difficulty when assessing subjective questions. This case has also been shown by the subsequent discussions regarding the approaches that have been used to assess subjective questions. Gyll and Ragland (2018) state that in a competency-based education, types of assessments determine the quality of this type of education because as Gyll and Ragland explain these types of assessments determine “an individual’s mastery of complex abilities” (p. 2). Therefore, the research problem according to the discussions so far includes the CAA systems inability to assess subjective questions. To clarify the definition of this research problem, more exploration is needed.

Gregor and Hevner (2013) suggest that a research problem definition should be constituted from the meta-requirement that Kuechler and Vaishnavi (2012) define as the behaviour of an artefact. Furthermore, Gregor and Hevner (2013) recommend the identification of the particular class of problem, which identifies the defined research problem. This class of problem is related to the classification of tasks including problems. The classification of tasks and problems are based on the ability of problem solvers to evaluate their capabilities/understandings for distinguishing relevant information (McKenney

& Keen, 1974). Furthermore, McKenney and Keen link the ability of applying these capabilities when considering the classification of tasks and problems. McKenney and Keen divide these problems into four classes. According to them,

> the fourth class of problem exists when both information and operations are unknown. In this situation, there is a conscious search for cues and a generation of explanatory concepts, together with the development of a method for manipulating the data thus organized. The development of new products is a typical research problem. (p. 86)

The other types of classes are discussed in McKenney and Keen (1974).

After the review and examination of the four classes, this study's research problem is associated with the fourth class that has been shown in the above quote. In this research, there is no data or information to uncover reasons of the inability of e-assessment, as well as no methods to uncover this information. There is also no available context to be studied and methods to be used to study this context to shed light regarding the inability of e-assessment subjective assessment. In this research, one of the research problems is: 'what are the problems with e-assessment when performing subjective assessment?' The consequence of identifying the problem class that the research problem associated with "helps with placing the work against prior literature and showing its contribution" (Gregor & Hevner, 2013, p. 349). After the identification of the research problem with the type of the class of problems, it is clear that this research problem falls under the 'improvement quadrant' goal. The goal of this quadrant will be discussed in section 1.2. The discussions thus far have been made about this research problem and the related topics indicate this research problem relevancy to a real-world practice.

1.2 Research purpose, scope and context

According to Gregor and Hevner (2013), one of the requirements in DSR project's content of the introduction section is to include the purpose and scope of the design science research. In a DSR, there are four quadrants: invention, improvement, exaptation and routine design. The goal of each quadrant is given in Gregor and Hevner (2013). Information systems researchers, when using the DSR paradigm, should identify their research with one of these quadrants. It has been declared in section 1.1 that this research is associated with the improvement quadrant.

> The goal of DSR in the improvement quadrant is to create better solutions in the form of more efficient and effective products, processes, services, technologies, or ideas.
>
> Researchers must contend with a known application context for which useful solution artefacts either do not exist or are clearly suboptimal. Researchers will draw from a deep understanding of the problem environment to build innovative artefacts as solutions to important problems. (Gregor & Hevner, 2013, p. 346)

The problem of assessing subjective mathematics questions when using e-assessment systems has been studied by the researchers in CAA for mathematics. Different approaches have been used to improve the ability of CAA systems. However, this problem is ongoing. The discussions that had been made in section 1.1 determined the link of this research problem with the fourth type of class of problems. The activities that are in this research scope, improvement quadrant, have been explained in the Gregor and Hevner's quote that has been displayed above. To achieve this research purpose, the fourth class of problem asks researchers to produce the knowledge that will be used in understanding the research problem, produce solutions as well as acquiring the ability to use these solutions to address the research problem.

In this research, a literature review will be conducted and the context which is student working will be analysed to extract the knowledge that has been mentioned above. The main purpose for analysing student working is to extract solution strategies as well as student errors in the working. The following qualitative and quantitative questions will be used as a guide when analysing the working. The purpose of analysing the student workings is not only limited to the answers these questions ask but to form the foundational knowledge that will be used to develop DSR artefacts. In summary, the purposes of this research are to produce and use this knowledge to improve the functionalities of CAA of mathematics when assessing objective assessment. Additionally, this knowledge will be used to produce DSR constructs and models artefacts that can be implemented to produce the prototype artefact that is capable of assessing solutions of an MSQ in the analysed student workings. This knowledge could be transferable to related disciplines. The following questions will be used to investigate student scripts.

Qualitative questions:

1) What types of solution strategies have been used to solve problems?

2) What types of errors have students made during solving problems?

Quantitative questions:

1) How many students have eliminated first, second and third elements sequentially in 3x3 matrix to create an upper triangular matrix in a Gaussian elimination?
2) What types of solution strategies have been used to eliminate the fourth, the seventh and the eighth elements from the matrix?
3) What is the frequency of this type of error across all scripts and where, in the student workings, does this error appear?

The research process that will be taken in this research will be explained in section 1.3.

1.3 Research process

In this research process, one of the initial tasks was to determine topics that will be reviewed to investigate the research problem further. The decision was made to select these topics: academic assessment, types of questions, functionalities of computer aided assessment systems, development methodologies, and other topics in related disciplines. Using the selected topics, the literature review was conducted. The review of literature reviews covering different topics as required has continued up to the end of this research. The outcome of the literature review in chapter two suggested the usage of the DSR paradigm to guide this research. The review was conducted to establish relationships between design science, natural and social sciences. Once the relationships between these fields have been established, a DSR framework will be created based on the outcome of the examination of selected well-known DSR frameworks. Various types of research philosophies were investigated to select the most relevant mythologies and methods. The selected methodologies and methods were used to analyse student scripts and create DSR artefacts.

Using the selected methods, student scripts containing student workings were analysed. The findings in the analyses of the workings were re-analysed to form categories and subcategories of solution strategies and student errors. A questionnaire was built and used in the quantitative analysis of student workings. The qualitative data of student workings was then quantified to perform the quantitative analysis of the workings. The findings from the quantitative and the qualitative analyses were used to form the knowledge that was used in creating DSR constructs and models artefacts. One of the models was implemented in building the prototype to be used to collect student workings. The artefacts were evaluated and demonstrated achieving this research purpose. The limitations in this research have been identified and reported. After claiming the research knowledge and suggesting recommendations for future works, this research is communicated in the structure that is presented in this thesis.

1.4 Research significance and contributions

In section 1.1, the importance of giving options to students when answering subjective questions has been discussed. This also allows assessors to measure students' comprehension of a domain and would give them opportunity to give better feedback. Subjective assessment evaluates "the conceptual grasping level of a candidate to how much the concepts are understood in a particular subject" (Dhokrat & Mahender, 2012, p. 14). A concept can be expressed well when students use their own expressions. According to Afamasaga-Fuata'i (2008), "students express the meanings in their own words perhaps with some illustrative examples … around the concepts" (p. 8). When comparing objective and subjective answers the answers that are given in a subjective assessment may possibly be satisfactory or correct (White, 2019).

McClarty and Gaertner (2015) identify that one of the requirements in an assessment design and implementation is to ensure that an instrument to be used is capable of performing its objective. Therefore, "program designers should work to clarify the links between the tasks students complete on an assessment and the competencies those tasks are designed to measure" (McClarty & Gaertner, 2015, p. 3). Subjective assessment is an instrument that is expected to produce the type of quality education, and the tasks students undertake when doing subjective assessment can be matched with the required competencies.

The contribution of this research is based on the outcome of the analyses of student workings that consists of subjective questions. Significant novel constructs have been formed from the findings of the analyses of these workings. The knowledge represented in these constructs is helpful in aiding program designers in understanding some important characteristics and/or actions students display/take when solving subjective problems. Furthermore, these constructs can be used in explaining methods or strategies that students use when solving questions with multi-step answers as well as possible errors that can be introduced. The knowledge that is displayed in these constructs can also be used to update teaching materials and for designing questions that aid in delivering authentic assessment. These types of questions contribute in obtaining lasting knowledge. The other significant contribution of this research is design knowledge.

One of the purposes of this research was creating DSR artefacts to implement method marking. Constructs, models and prototype DSR artefacts were created. These artefacts contain significant contributable design knowledge. The models artefacts can be used in explaining problems in the investigated domain as well as producing solutions. Moreover, these models contain knowledge that is transferable to other similar contexts or disciplines. The prototype artefact that implemented some of the models also show design knowledge. The approaches that haven been devised in implementing the method marking is design knowledge that can be claimed as significant contribution. The design knowledge, which is produced in this research, can be practically implemented in developing CAA systems. The other significant knowledge, which has been uncovered in this research, is found in this study's research process.

The methods that have been used to analyse student workings and create DSR artefacts are novel, and as such can be claimed as significant contributable method knowledge. During this research, specific methods have been developed to analyse student workings. The approach that has been used to analyse these workings has been general, although the specific context has been analysed. Section 5.3, which discusses categorisation of the outcome of the analysis of student workings, details this approach. The other method knowledge contribution is the approach that has been used in designing the questionnaire. This questionnaire is a significant instrument that can be used in contexts that entail the investigation of multi-step questions and/or multi-step problem.

1.5 Structure of this thesis

In chapter two, the research process that will be used to conduct the literature review will be explained. The 'qualitative content analysis' approach will be explained to show how this literature review will utilise this approach to extract relevant information from selected publications. Different types of assessments in education will be identified. Types of questions in assessment in addition to the type of knowledge these questions produce will be discussed. The assessment tools which are computer-aided assessment (CAA) systems for mathematics will be investigated to determine types of questions they are able to assess and how they assess them. Furthermore, the development methods of these CAA systems will be reviewed. The findings from the review will be discussed and emerging gaps will be investigated further. To perform this investigation and other research activities, this study's research process will be explored in chapter three.

The discussion in chapter three starts with examining relationships between the design science and other science branches. The chapter continues to the next stage which is reviewing selected DSR frameworks to construct one that is appropriate for this research. In this chapter, research philosophies will be examined to select the most relevant ontological and epistemological stances concerning the scripts and DSR artefacts. Using these selected stances, methodologies will be selected. The selected methodologies design will be examined in chapter four to select the most relevant qualitative and quantitative research designs.

At the outset of chapter four, different types of qualitative design will be reviewed to select the most appropriate design to analyse the student workings in the scripts. The sampling process regarding the student scripts will be described and the relevant sampling process will be adapted. Using the inductive approach, the analysis of student workings will be performed. After the qualitative analysis of the workings, different types of quantitative design will be examined to select the most appropriate design for the quantitative analysis of the scripts. The survey design will be examined, and the required questionnaire instrument issues will be discussed. The results of both qualitative and quantitative analyses of student scripts will be presented and discussed in chapter five.

Chapter five starts with discussing attributes that verify the acceptability of the analyses of student workings process and the outcome of this process. The process that is used to categorise the findings in the qualitative analysis of the student workings will be discussed. The categorised findings in conjunction with their meanings will be presented and discussed. Following the presentation of the result of the qualitative phase, the quantitative analysis will be reported and discussed. The relationships between categories/subcategories will be evaluated and models will be created. Review of the scientific theory will be made to discuss the theorisation of these categories (constructs) and models. Using these constructs and models, DSR artefacts will be created in chapter six.

At the start of chapter six, the characteristics of design science research artefacts will be identified. Some terminologies in DSR in addition to their relationships will be discussed. The importance of constructs, models, methods and instantiation artefacts as the outputs of a DSR will be described. Constructs and models artefacts will be created from different sources especially from the knowledge that was created from the analyses of scripts. Specific features from the models' artefact will be demonstrated by creating the methods artefact. The prototype artefact will also be created to show the possibility of implementing a model artefact. Finally, the evaluation methods of the prototype artefact and methods artefact, and the result of the evaluation will be discussed. Furthermore, the issues of evaluation will be explored in chapter seven.

The focus in chapter seven will be on the evaluation purposes, methods of evaluation and validation of the research output. It also discusses the importance and scope of evaluation processes in research. Additionally, it discusses the methods of evaluation that have been used in this research and the attributes of this research's artefacts. The completion of the evaluation processes leads this research process towards a final discussion and conclusion.

Chapter eight starts with summarising activities in this research. Following the summary, the findings in the analyses of student scripts will be discussed. In addition, the findings related to the development of this research artefacts are discussed. This chapter presents the knowledge that is contributed and shows how the contributed knowledge was acquired. Finally, the chapter lists the limitations in this research and ends by presenting the future directions of this research.

1.6 Chapter summary

In this chapter, a review about the research background was conducted to uncover this research problem. The emerged research problem was investigated further and finally defined. The examination of classes of problems helped to identify the class that is related to this research problem. The identification of the research problem and class helped determine the scope of this research.

The research purposes and the scope of this research were explained and the context to be studied was determined. To achieve the research purposes, the 'research process' was explained. The significance and contribution of this research have also been identified.

In the end, the structure of this thesis was presented including the brief introduction of each chapter.

The next chapter includes an expanded discussion of the research background and the research problem.

Chapter 2 Literature review

The introduction chapter has defined the thesis focus identifying the context to be studied and declaring the research project aims to be achieved. The chapter has also categorised the problem class that enables the researcher to delineate the research problem. The chapter has, further, highlighted the purpose and scope of the research, the research process and the significance of the research. The purposes and approaches that will be used to conduct this literature review will be presented after the definitions of the following terminologies.

Some of the terms and/or concepts, which are not familiar, appear in different sections of this chapter; their definitions are given next.

- "Ill-defined problems are those problems that do not have clear goals, solution paths, or expected solution" (Arifin, Putri, Hartono, & Susanti, 2017, p. 2); the ill-defined, 'wicked' and 'complex problems' concepts share similar characteristics;
- Mal-rules "are wrong paths producing errors which can be classified as mechanical, conceptual, procedural or by application, Schechter (1994)" (Hanson, 2011, p. 22).
- 'Follow through' error term is used when a student working that contain error(s) is penalised once and the subsequent correct workings are exempted from the penalty, although they use the previous step's incorrect answer (Pead, 2010).
- 'Method marking' concept is the notion that awarding learners' workings is based on the technique(s) they use in solving questions that require the demonstration of solution strategies to produce a final answer that might be correct or incorrect (Pead, 2010).
- 'Partial credit' term is applied for the award that is given "based on "working that could lead to a correct answer" or other marks that can be awarded without a correct answer to the subtask" (Pead, 2010, p. 111).

Before examining the intention of this chapter, it is important to restate the aims of this research project in order to link the targets of this chapter to the goals of the research.

The purposes of the thesis is to (1) analyse a phenomenon that produces the knowledge that assists in understanding the challenges that inhibit developing and evaluating e-assessment for mathematics, also called computer aided assessment (CAA) of mathematics and (2) to create the DSR artefacts that will contribute solutions to overcome or manage the challenges that will be described in this chapter. In order to comprehend such challenges and form the relevant knowledge, it is important to explore the contexts that CAA systems are employed in and their utilities. Therefore, in this chapter, the review will cover: (1) the role of effective academic assessments in education, (2) the utilities of the current e-assessment concerning the assessment of solutions to mathematics' questions, (3) the methodologies that have been used to develop CAA of mathematics and (4) the natures of mal-rules, wicked or complex problems and complex problems solving. These subjects will be thoroughly investigated in the subsequent sections and the findings will be discussed and conclusion will be made at the end of the chapter. *The structure of this chapter* is as follows.

Section 2.1 describes the literature review methods, review type and the importance of using the research process. The section, likewise, elaborates the importance of taking philosophical assumptions for better comprehension of the literature to be reviewed and to be able to select relevant research methods to conduct the literature review. A brief explanation of the selected method: qualitative content analysis is presented in section 2.2. In this section, techniques to determine the number of publications to be analysed and to search for the relevant literatures are explained. Furthermore, the section details how to manage the collected literatures and analyse them. Finally, the section reports how data from each publication is examined and synthesised and grouped to describe the topics. Using the content analysis method, section 2.3 and its subsections focus on the assessment topics. The academic assessment and its different types of assessments are identified and described. The purposes of various types of assessment in extracting information for different purposes are also covered. Additionally, the type of assessment questions such as objective and subjective and their roles in revealing different levels of knowledge are defined. The section concludes with explaining the role of authentic assessment in helping students to receive meaningful knowledge.

The acquired knowledge requires relevant assessment tools. In this research the subject matter is the CAA tools, which section 2.4 and its subsections explore. The brief descriptions of CAA systems are given. The criteria to assess student workings such as method marking, partial credit and follow through concepts are, likewise, described. The importance of the suitability of an assessment tool to assess a subject that it claims to evaluate is also addressed. The section also includes the explanations of the type of questions that CAA systems can assess efficiently as well as the limitation they have to assess subjective questions. In the same section three different designs to assess subjective questions are explored. The section ends by suggesting the exploration of current CAA systems development methodologies.

In section 2.5 the CAA systems development methods are reviewed and examined. In this section, the ontological and epistemological assumptions that have been taken in the review have been identified. The reviewed articles discuss methods used to develop the components of CAA systems that include the determination of assessment purposes, creation of new assessment questions or searching for existing question banks, question authoring and randomisation techniques; this section also discusses the methods used to collect students' answers, assessing them and giving feedbacks to students. The CAA system development is generally based on objective questions due to the complexity of assessing subjective questions.

The challenges that are presented by complex problems are explained in section 2.6. This section begins with the discussion of the issues of identifying errors that are made by students. It further explains the causes of these errors and the identification of mal-rules. The relations between knowledge levels and student errors are also explained. The similarity between the difficulty of identifying some student errors and the nature of complex problem solving is realised. This section, likewise, describes the general concepts of a problem, the features that qualify a problem to be complex and the characteristics of complex problem solving. Using these characteristics, the associations between CAA systems subjective assessment and complex problem solving are established. This section identifies the complex- problem solving structure; using the structure, the subjects that help in solving complex problems are detailed. The links between knowledge accumulation and solving complex problem are described. This section also reflects on the purpose of DSR in relation to solving wicked problems in IS research. This section concludes on showing the connections between CAA systems and DSR IS artefacts.

2.1 Literature Review Process

In order to achieve the goals of this review, different literature review techniques have been explored to select the most appropriate literature review method. The researcher has also investigated the type of review that will achieve the review's purpose. The selected type of literature review was "review for understanding" (Rowe, 2014, p. 244). According to Row, the purpose of this kind of review "aims at identifying key findings, problems, and research thrusts and paths to solve them" (p. 244); and the relevant review method that has been selected was the systematic review, which is "characterised by being objective, systematic, transparent and replicable" (Siddaway, 2014, p. 1). The literatures have been collected and analysed using the selected review method to achieve the purposes of the review.

Research processes are required to investigate research problems in primary research as well as to conduct literature reviews, which is also called secondary research. To show the importance of the secondary research, similar research process that is applied in the primary research is used; because the outcomes of the secondary research serve as the catalyst for "the primary source of research questions" (Guest, Namey, & Mitchell, 2012, p. 17). The researchers' familiarity with the research philosophies which encompass ontology, epistemology, methodology and research methods help understanding the contents of the reviewed research literatures and their research findings and/or outcomes. The process of selecting the relevant research methods and research philosophies are detailed in section 3.3.1. Based on the explanations that are given in section 3.3.1, the researcher has taken the assumption that the reviewed literatures are identifiable real-world objects revealing valid knowledge. The assumptions that have been taken are the critical realism stances which are based on the philosophies of interpretivism ontology and of constructivist epistemology. Using these assumptions, the inductive qualitative content analysis approach, which is explained in section 4.1, has been chosen to analyse and critically review the publications.

2.2 Content Analysis Process

The general steps that are required in a qualitative content analysis method are shown in Figure 4.1. The figure shows (1) sampling, (2) collecting and managing raw data, (3) preparing raw data, (4) organising the analysed data and (5) reporting the result of the analysis. In this literature review, there is no exact number of publications to be determined to accomplish the purposes of this review. The only measure, for the number of publications to fulfil the purpose of this review, is the amount and the quality of the publications that can be justifiably judged to produce sufficient knowledge that is adequate to build foundations for beginning the primary research. The following section includes a brief description of the qualitative content analysis method that is applied to conduct this literature review.

2.2.1 Content analysis phases

This section includes the phases of preparing, collecting, managing and organising literatures; it further includes coding and forming concepts. In Figure 4.1 the preparation phase is shown at step three. In this phase, the objects to be analysed, the extent of analysis and the unit of analysis are identified. To identify the publications to be reviewed, initially many different terminologies, in the current research and related disciplines, have been used to find the associated key words and concepts. The process was iterative and took time to distinguish the applicable concepts, themes and topics to be studied. This procedure had helped in establishing what to analyse, resolving the question of the unit of analysis for conducting this literature review. According to Randolph (2009) the units of analysis in “a literature review are the articles that are reviewed” (p. 4). More detail about a unit of analysis is given in section 4.4.1.3. The actions thus far have been taken have resolved the issue of the criteria of the inclusion of publications based on the subjects they cover. The next step is to collect and manage the publications to be reviewed.

The collected publications were classified into relevant and semi relevant groups. The relevant and semi-relevant literatures were subdivided based on the topics they covered. The next phase was the organising phase. Organising, in the content analysis method, involves evaluating the studied objects, which are the publications that have been selected for reviews, in the current context. The following steps have been employed in the publications’ evaluation process:

1. Read the literatures for understanding.
2. Determine the topics or the themes the literature cover.
3. Add a literature to the category/theme it belongs to.

4. Number the literature for the identification purpose.
5. Reread the literature again for further understanding and clarification.
6. Highlight the relevant terminologies, definitions, and authors' views, research findings and methods and add comments (reflections) to the highlighted part of the literatures.
7. Summarise the literature, if required.
8. Save commented literatures and summaries electronically and in hard copy.

These steps have led to the creation of succinct pertinent manageable knowledge that will be used in the next phase as the components to construct the new knowledge. The next phase performs grouping, coding and forming concepts or themes. More details of these actions are given in section 4.4.4.

In this grouping and concepts forming phase, the manageable knowledge that had been formed in the previous phase has been synthesised. This knowledge that comes from different publications has been examined critically and the relationships between the data in each publication are the factors that have determined the formation of the synthesis. During publications identification and examination process, the interpretation process helped the understanding of the investigated topics. Finally, the synthesised knowledge is used to form concepts or constructs that have been used to explain the topics and/or the context of the topics. The content analysis process that has been briefly explained in this section is implemented in the subsequent sections.

2.3 Academic Assessment

Assessment is a procedure that involves students and teachers as the main stakeholders in which teachers judge students' works and give them a feedback based on the outcomes of the judgement (Sangwin, 2013); and the given feedback is purposeful determining the future step(s) that influence(s) directions taken by the assessment's stakeholders (Huba & Freed, 2000; Kellaghan & Greaney, 2001). Assessment is viewed as the main element in teaching and learning processes, and for students it is the most important part of their learning; assessment is seen as the main activity that influences their approach to learning (Price, Carroll, O'Donovan, & Rust, 2011; Sangwin, 2012). In a case study that was conducted to reveal the influence of assessment on students, it was found "not only did the assessment generate enough learning activity, it generated appropriate learning activity" (Gibbs, 1999, p. 45). To reveal different purposes, academic sectors use several forms of assessments.

2.3.1 Types of Assessment

Formative, summative, diagnostic and evaluative are the most used types of assessment in academic sectors. Formative assessment is used to check the current students' status, where they are heading and how to reach their targets (Wiliam, 2011). It also directs the teaching activities (Gears, 2005) and contributes to the students' lasting knowledge (Shepard, 2005). It was found that students who had been given successful formative assessment tasks, showed high achievement in summative assessment tasks (Balcombe, Brennan, & Everiss, 2011; Black & Wiliam, 2005; Gears, 2005).

Summative assessment is designed to measure the achievements of learners for the purpose of grading their works and is normally performed at the completion of a course (Cook & Jenkins, 2010). Diagnostic assessment is used to check students' prior knowledge and skills of a particular domain, and to gather from the assessment results the perceptions and misconceptions that are shown by students (Ketterlin-Geller & Yovanoff, 2009). Such assessment is performed "at (or even before) the beginning of the teaching program" (Cook & Jenkins, 2010, p. 8). Sangwin (2013) also states that diagnostic assessment is used to uncover the level of previous learning for the purpose of aiding students to focus on particular topics or knowledge. On the other hand, an evaluative assessment is utilised to gauge the effectiveness of educating or assessing students (Sangwin, 2013); and it is likewise "used as an indicator of individual school/teacher performance in terms of education quality and the effective application of policies and practice" (Luckin, Clark, Avramides, Hunter, & Oliver, 2017, p. 86). These types of assessment measure specific or general knowledge that students should display or knowledge about resources; in the case of evaluative assessment, to deduce information that assists stakeholders to take better and more appropriate decisions. The significance of the academic assessment in education leads to the necessity to explore factors and tools that assists in applying authentic assessment that is able to test or produce profound and applicable knowledge.

2.3.2 Assessment authenticity, questions' roles and questions' types

Authentic assessment measures important skills that are regarded as "intellectual quality" (Palm, 2008, p. 7) that contributes to the assessed students' achievements. It is important to ensure that students understanding of a domain is fulfilled by exploring students' knowledge through assessment that is viewed as "the bridge between teaching and learning" (Wiliam, 2013, p. 15). Various types of assessment methods are utilised in academia. Sim, Holifield, and Brown (2004) claim that fifty different techniques have been used in higher education to manage academic assessment and this number is increasing. Among these techniques, according to Graham (cited in Sim et al, 2004), exams and essays were the most prevalent. Most written exams consist of questions and the function of questions in advancing learning is significant. In all theories of learning, the integral element is asking students questions (Sangwin, 2006). For example, the actual learning of mathematics is realised when learners' participation in the learning becomes active process (Sangwin, 2006). Posing questions to learners helps to promote this active process. Another example that emphasises the role of questions in learning is demonstrated by Fernandez & Makoto (as cited in Wiliam, 2005). Makoto (as cited in Wiliam, 2005) stated that teachers, in Japan, spend significant time in developing questions to measure their teaching success. The two examples show the roles of questions in learning and teaching efforts.

The standard question's types in academic assessments are subjective and objective. According to Suskie (2010) "subjective assessments yield many possible answers of varying quality and require professional judgement to score" (p.33). Here Suskie is referring to the type of assessment that uses subjective type questions. Objective questions or assessments exclude assessors or examiners from influencing the outcomes of the assessments, while restricting the assessed students to select the answers from the given options (Sangwin, 2013). Examples of objective questions' type are multiple-choice and multiple-response (Lilley, Barker, & Maia, 2002). Objective type questions, in multiple-choice format, are also given different names; for example, selected response questions (SRQ). The subjective questions' type examples include "short answer, text match, explanation, numeric and calculated (generated data, parameter dependent) questions." (Hu & Xia, 2010, p. 250). Both types of questions have different influences on the knowledge they produce or measure.

There are various opinions about the quality of SRQ assessment form. One view is related to this type of questions' inability to produce high level knowledge. Veloski, Rabinowitz, Robeson, and Young (1999) point out that presenting physicians with the questions based on five different choices to select one correct answer cannot be compared with the actual scenarios involving human patients. Furthermore, Veloski et al. (1999) claim that SRQ type of questions assess trivial knowledge at the expense of fundamental knowledge. Simkin and Kuechler (2005) conducted a knowledge-level analysis using both SRQ and constructed response question (CRQ) tests and questions bank belonging to a basic computer language programming domain using Bloom's taxonomy of learning goals for the domain. The aim was to analyse the data to discover the level of knowledge the two types of questions could assess. The result of the analysis revealed that SRQ questions assess different levels of knowledge showing the nature of inhomogeneous characteristics. When comparing the two formats of questions, SRQ less able to test the highest level of knowledge that CRQs were able to assess. Hence, the acknowledgement of the link between the assessment questions and the knowledge that is expected from the assessment outcome is important.

As mentioned above, objective assessments produce surface or trivial knowledge; on the contrary, subjective assessment tests high order knowledge or produce thoughtful knowledge. The difference in the outcomes of the two types of assessment thus have impact on the stakeholders, especially students. Bloxham and Boyd (2007) assert that genuine assessment results guide students towards deep learning leading them to attain lifelong knowledge. Students use deep and surface methods to approach learning. In the deep approach, students relate information and ideas to each other and also to their own experiences, and look for patterns, principles and meanings in the texts; but in the surface approach, the goals are just to meet tasks' requirements (Bloxham & Boyd, 2007). The high-quality knowledge creation is not only the consequence of using good questions in assessment but also involves multiple factors; among them is selecting an appropriate tool which determines how and what to assess. The focal point in the next section is the CAA systems tool.

2.4 Computer Aided Assessment

Various online assessment technologies, with varying names, are in use. E-assessment is one of these names and it is defined as: "E-Assessment is the use of information technology in conducting assessment" (Singh & de Villiers, 2017, p. 164) that involves "the presentation of assessment activity, and the recording of responses" (JISC, 2007, p. 6) using ICT assessment tools. E-assessment involves the planning of assessment questions, storing them in media, displaying to users, collecting users' responses, giving feedback and keeping the entire assessment process records (Cook & Jenkins, 2010). The assessment tasks are "mostly or completely marked automatically using Information and Communications Technology" (Cook & Jenkins, 2010, p. 2). A form of e-assessment tool utilised in assessing mathematics questions is CAA system; and a focus of this literature is revealing the features of e-assessment for mathematics. Note that the discussion about CAA systems in this thesis is focused on CAA for mathematics. Therefore, if the phrase 'for mathematics' is dropped, the abbreviation CAA applies for CAA for mathematics. These types of CAA systems need assessment criteria to assess student workings.

2.4.1 Assessment Criteria and Method Marking

Assessing student workings for awarding marks requires marking guides that detail marking criteria. These criteria, in the context of e-assessment and/or CAA systems as assessment tool, often include the discussions of method marking, follow through error and partial credit phrases. The meanings of these phrases of concepts have been given at the start of the chapter. Pead (2010) has stated two things that are the main concerns of assessing mathematical subjective questions when applying 'method marking' concept, using e-assessment tools: (1) the assessment work does not only require assessing workings in the student workings when assigning partial credit, but it also involves searching for correct and/or incorrect works in the overall workings to decide whether to assign partial credit or not. (2) The second issue is students' inputs of their workings, which as well includes solution strategies. This second issue is the interface matter that can be used to guide students managing their inputs. For example, boxes to collect a variable elimination method, one box could be to input a variable name or symbol and the other box to input the elimination formula or method. Hence, the task of awarding marks to student workings using e- assessment tools does not only involve setting marking criteria but also the ability to accurately implement these criteria by using techniques, such as method marking.

2.4.2 How do CAA systems assess mathematical answers

The assessment tools should be relevant for the subject they assess to achieve successful assessment outcomes. For example, the characteristics of assessment in mathematics differs from other disciplines (Palm, 2008). Mathematics assessment nature is mostly subjective. As explained in section 2.3.2, subjective and objective questions types are prevalent in academic assessment. Computer aided assessment systems, generally, assess objective questions (Sangwin, 2013) and these systems face challenges to assess “all possibilities of students’ responses, as these sometimes may be unlimited” (Adesina, Stone, Batmaz, & Jones, 2015, p. 95). Therefore, “the focus of many CAA systems has been the final answer or product of assessment tasks” (Adesina, Stone, Batmaz, & Jones, 2014, p. 1). Sangwin (2015a) and Paterson (as cited in Sim et al., 2014) indicate the impracticality of CAA testing higher order knowledge; while Lawson (2012) and Mackenzie (as cited in Jordan, 2014) accept that CAA is not a complete panacea in assessing multi-steps questions. The acknowledgement of the difficulty of assessing subjective questions satisfactorily has led researchers to devise various approaches to improve the capability of these systems. The three general approaches that have been used to assess the subjective questions are comparing expressions, dividing multi-steps questions into sub-steps and trapping and using common errors and misconceptions in the form of feedback.

The first design is expressions comparison. In the expressions comparison design, the match between students’ supplied expressions and the questions setters’ expressions is performed to check for the correctness of the students’ answers using a computer algebra system (CAS). Sangwin et al. (2009) explain that the CAA systems that are supported by

CAS check for the properties by assessing the answer, which is in the form of an expression. The correctness of the properties in the student's answer determines the awarding of partial or full marks. As for the CAA systems that do not use CAS, they have used their own techniques to perform expressions comparisons and awarding marks.

The second design is splitting multi-steps questions. In this design, the method that has been applied is to split multi-steps questions into sub-step questions with one step and awarding marks each step separately. The Project for Assessments in Scotland using Information Technology (PASS-IT) project used this method without applying CAS (Ashton et al., 2006, p. 98); in a study, the PASS-IT project used two different ways to tackle the assessment of subjective questions. After breaking long questions into steps, two options were given to assess student workings. One option was allowing students to seek steps display and awarding marks only for implementing the strategy given in the steps. The second option was assessing student workings without displaying steps. In both options the system accepted mathematical expressions from students and compared them with the questions setters' expressions. Beevers et al. (1999) used similar techniques as PASS-IT project; however, they did not penalise students for marks for seeking steps.

The third design involves assessing student workings by trapping common errors, such as arithmetic and algebraic, and displaying them as feedback and giving opportunity to students to correct the errors and resubmit their works.

Numerous types of CAA systems for mathematics have been developed using various techniques and it is impossible to discuss each design. However, it is worth denoting that the general approaches that have been used to assess MSQ (subjective questions) are very similar. Despite the claim that these approaches assess MSQ solutions, Sangwin et al. (2009) explain that the feedback produced, as the results of comparing expressions, do not contribute to the improvement of students' mathematical understanding: because the information to give the feedback is based only on the final answers that do not include student workings. This section has revealed what CAA systems can and cannot assess when they are used as assessment tool for mathematics. The artificial phenomena, which CAA is one of them, “can be both created and studied, and that scientists can contribute to each of these activities” (March & Smith, 1995, p. 253). The review of design and development methodologies and methods of CAA systems or generally e-assessment systems would help in finding and understanding constraints that limit the construction of CAA artefacts. Furthermore, the knowledge that is deduced from the outcome of this review adds some knowledge that might be helpful in developing artefacts. Furthermore, the knowledge that is deduced from the outcome of this review adds some knowledge that might be helpful in developing artefacts. This knowledge could contribute to the improvement of the assessment of subjective type questions and/or could also enhance the quality of the knowledge that emanates from objective type assessment.

2.5 CAA Development methodologies and methods

The interests of researchers regarding IS/DSR artefacts entities will be explained in section 3.1; and similarly, in section 3.3.4, the ontological and epistemological stances regarding IS and/or DSR will be explained. In the current literature review of CAA systems, many of the accessed articles have not discussed the issues of research philosophies regarding the CAA systems. However, one of the articles has highlighted the subject of ontological and epistemological assumptions in regard to CAA systems implicitly. Rønning (2017) explains how knowledge is “actively constructed by the learner in interaction with other learners, the teacher and various learning resources provided” (p. 96). The CAA systems are one of the learning resources that students interact with to obtain knowledge. This philosophical assumption shows that CAA systems (artefacts) are the real-world objects. This philosophical assumption will be adapted in the review of the development of the components of CAA systems and how they are used to build the systems.

When developing a CAA for mathematics system, excluding CAS, the main focal points are (1) preparing questions, (2) authoring questions, (3) setting assessment criteria, (4) collecting students' responses, (5) assessing students' responses and (6) giving assessment feedbacks. The assessment of students' responses and feedback methods and the implementation of the methods have been identified and explained in section 2.4.2. Some of the main assessment criteria, in CAA system, such as method marking issue have also been explained in section 2.4.1. In this section the review will concentrate on uncovering the methodologies that have been employed in resolving the subjects' one, two and four, which have been listed above; the section likewise adds some details to the subject five, which is the issue of assessing students' responses. The researcher believes that the activities in the development of a CAA system for mathematics can be, mainly, divided into two blocks. The first block, which is shown in Figure 2.1 could be labelled as the 'prior assessment delivery' (PSD1) and the second block as the 'post assessment delivery' (PSD2).

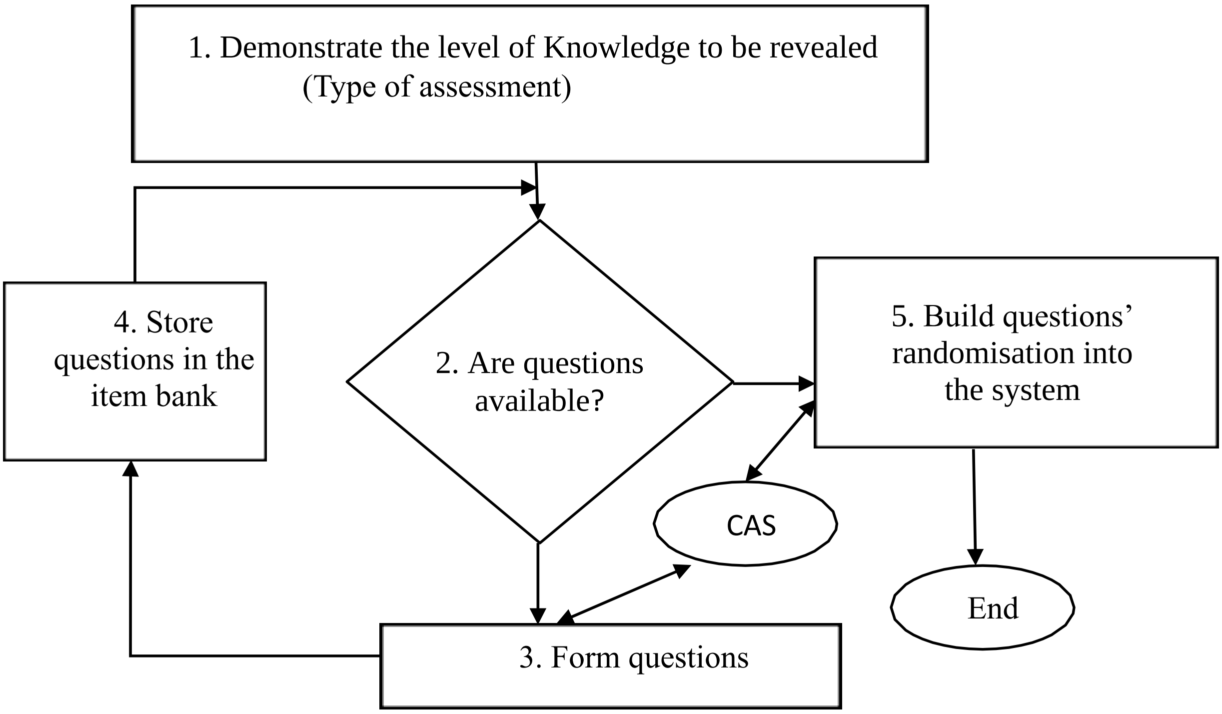


Figure 2.1 Prior assessment block

The first stage in the PSD1 block is determining the level of knowledge to be tested. At this stage establishing the type of assessment, whether subjective, objective or both, is necessary. In the guidelines of writing effective questions for CAA systems, Greenhow (2015) explains it is important to know the purpose of assessment; because each type of assessment is tailored for a different purpose and, therefore, the assessment questions should be designed accordingly. Hence, searching for the sources and/or formation of questions that achieve the assessment purpose is fundamental; this activity is represented by the stages two and three in Figure 2.1. The sources of the questions could be item banks, or newly made or transformed from paper-based assessment into computer-based assessment. According to Chilivumbo (2015) this could be by:

- selecting questions from different books but pertaining to the content … lectured in.
- creating Short answer questions and Essay-type questions.
- using materials used for teaching, course outline objectives, and other end of book exercises.

Other methods used to create questions include:

- textbook question banks (hard copies).
- past assessments (reuse of past questions).
- past assessments (modifying past questions).
- general Internet searches. (p.29).

The word item refers to the word question; therefore, the two words will be used interchangeably. Item banks are defined as “collections of questions, often produced collaboratively across a subject domain” (Conole & Warburton, 2005, p. 22). Sclater (as cited in Conole & Warburton, 2005) believes that the question banks are one of the driving factors in the development of CAA. If the item banks are empty, questions should be formed and transferred into the banks. Copy like translation of paper-based assessment to computer-based assessment is unlikely. One of the findings from the literature review, spanning over a decade, vis-à-vis designing of questions for CAA development, reveals that the direct conversion of a paper-based assessment to a computer-based and/or online assessment is inapplicable (Conole & Warburton, 2005). Commenting on this issue Sangwin and Köcher (2016) have stated that "to start with an existing examination format and merely translate questions into a new format without regard for the underlying educational construct they are seeking to test might seem incongruous" (p. 216). The rationale for "understanding these issues is important for developing strategies for item development as well as to produce guidelines for developing appropriate administrative procedures or statistically adjusting item parameters" (Conole & Warburton, 2005, p. 21). However, irrespective of the sources of the assessment questions for CAA systems, questions should be integrated into the system to be used in the assessment. The type of questions that CAA system can assess is, generally, objective and this fact has been mentioned in section 2.4.2. The attention, here, is to find the methods that are used to form objective questions such as multiple-choice questions (MCQ). MCQ are consist of "stem, key, equations, figures, feedback" (Greenhow, 2015, p. 3); Greenhow emphasises that in addition to these components, MCQ should include parameters. The components of MCQ require careful planning to give credibility and validity to the knowledge that is produced by MCQ. Greenhow suggests designing the objective questions using means such as asking students to recognise "where a mistake might lie (if any) or a multi-choice question (MCQ) might offer well-chosen (mis)interpretations for a student to select" (p. 2).

Another method that has been implemented is using mal-rules; Walker, Gwynllyw, and Henderson (2015) have used mal-rules that have been discovered in the previous research as distractors in forming MCQ. More information about the nature of mal-rule is in section 2.6. The detail of how CAA systems manage multi-steps question is given in section 2.4.2.

The above-mentioned methods and similar suggestions, if any, are able to resolve the issue of the formation of objective questions. The other issue that raises interests in CAA system development is authoring questions.

Teachers need means to help them write questions that test students to display specific knowledge. Greenhow (2015) sees the difficultly of teachers using their own only efforts to author questions due to the high demands, which are "pedagogic and technical levels" (p. 4). Some of the CAA systems did not provide CAS utilities forcing teachers to code their own computer algebra code for each question (Sangwin, 2015b). However, fortunately the CAS utilities are available in some CAA systems such as STACK and Maple T.A to help teachers author questions; some commercial CAA systems allow teachers to author their own questions without relying on the questions that are in the systems. This status is shown in Figure 2.1 by the CAS utilities connections to both 'form questions' and 'questions' randomisation' blocks, which will be explained later. The CAA systems require a method(s) to be able to use questions having the same structure, for example $a^2 + b^2 =$ 13, to be used multiple times. The questions' randomisation technique, which is implemented by CAS, addresses this issue using parameters which "usually generate different numbers, and can also generate random words or question scenarios and are certainly not limited to the assessment of mathematical content" (Greenhow, 2015, p. 3). Using the parameters technique, the above example could be displayed in different ways such as 2a2 + 3b2 = -21, or a2 - b2 = 5, etc. Therefore, parameters can assist in generating "many thousands or millions of realizations" (Greenhow, 2015, p. 3). For example, Maple T.A system can create multiple questions from one template by changing parameters of the template questions (MAPLESOFT, 2016). However, being able to present students with such huge numbers of 'different' questions does not mean the questions are completely different; only the parameters are different. It could be stated that "for most problems, it would be possible to create a general solution to the problem, and then in each case just substitute the particular values of the parameters into the answer "(Rønning, 2017, p. 99). The explanations thus far have been presented, generally, cover the CAA system's development activities that have been shown in Figure 2.1; in other words the activities in PSD1 block are complete and the questions are ready to be used in a CAA system assessment.

The PSD2 block that is shown in Figure 2.2 allows the delivery of the questions to the students, collect responses from students, assess these responses and return feedbacks. As the request is made, the system applies the randomisation process and displays the question. Figure 2.2 shows the CAS part of the block connections to the three stages in the block. The CAS utilisation is present at the stages that include questions authoring, as mentioned above, delivering questions to students, reading their responses and assessing them. The activities in the boxes labelled 2, 4 and 5 in the Figure 2.2 indicate these issues. When the system reads a student's response, stage 4, CAS checks for the right syntactic (Sangwin, 2015b) to accept or reject the student's response. Therefore, students are required to know how to use CAS.

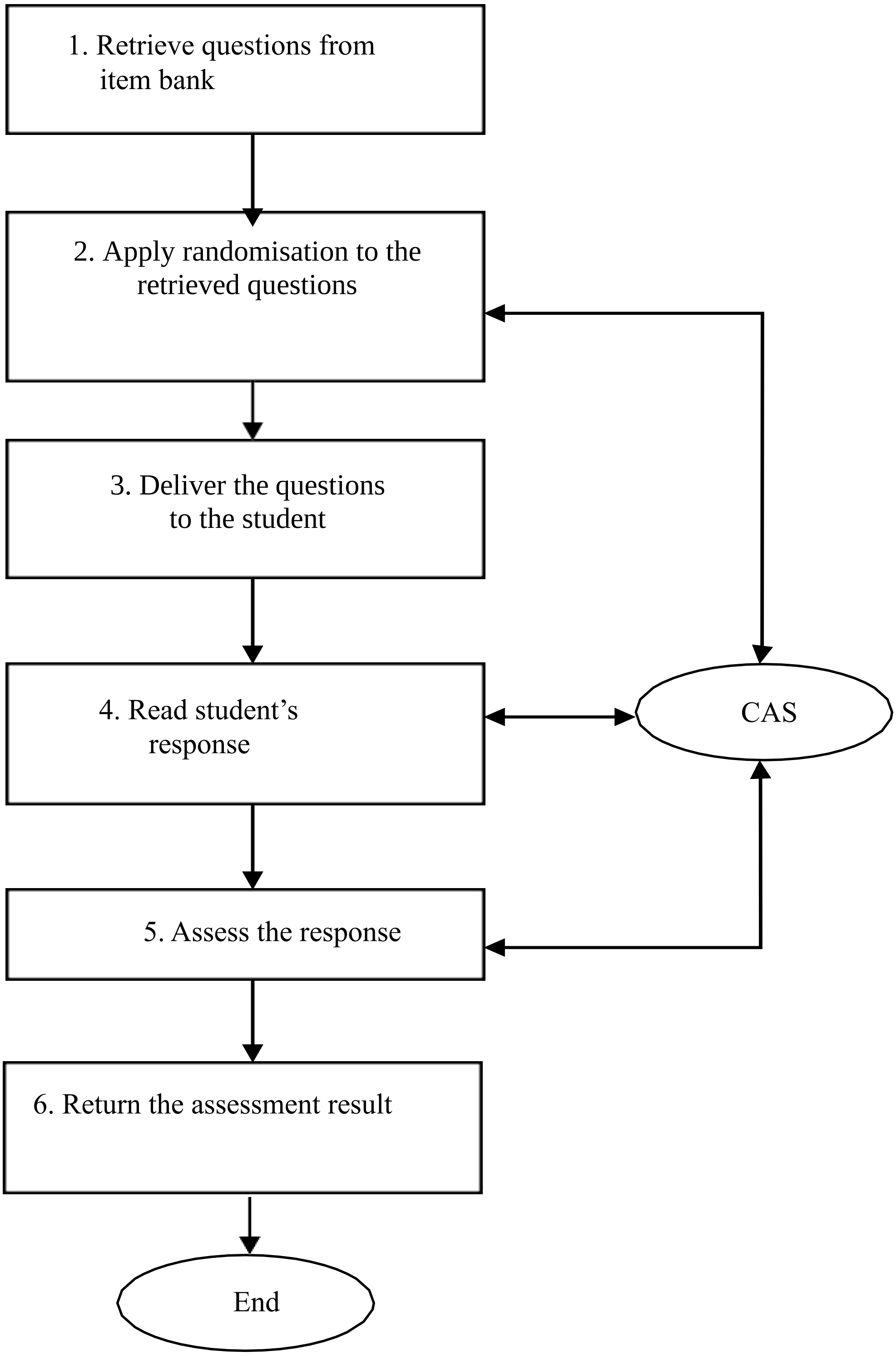


Figure 2.2 Post assessment delivery block (PSD2)

The student's answer could be in MCQ form or filling blank with the answer that could be numerical value or algebraic expression (Rønning, 2017). The algebraic expression is entered using equation editor that is provided by Maple T.A system (Rønning, 2017).

When assessing students' works, CAS performs the assessment process as explained in section 2.4.2. The CAS system provides, for example, the 'simplify' command to look for the equivalency of expressions. This command returns true if the two expressions are equal or false otherwise. The way the simplify command is implemented is dependent upon the CAS system that have been developed by different developers Sangwin (2015b). The simplify command example shows varieties of CAA systems with varying capacities. However, the general CAA systems development methods are summarised in the activities that are displayed in the two figures. The challenges that are related to the development of CAA systems will be the focus of the next section.

2.6 Complex problems and complex problem solutions

The limitations of CAA systems assessing full student workings have been identified in section 2.4.2. One of the limitations was, the unbounded or unknown responses and natures of students' workings (Adesina et al., 2015) as mentioned in section 2.4.2.

Likewise, it is impossible to predict the type and number of errors in student workings. Payne and Squibb (1990) conducted an empirical study to uncover elementary algebra errors. Among their findings were varying students' abilities displaying different outcomes. Some outcomes have showed errors that can be identified and classified as a common error of a particular type. According to Hanson (2011) a common error is caused when students are unable to produce correct answers because of the absence of some knowledge and the presence of the breakpoints that prevent them from going forward. Payne and Squibb state that common errors can be used to classify mal-rules. Greenhow (2015) labels mal-rules as "specific mistakes that are likely" (p. 5) and Greenhow considers them as errors that occur often at a basic stage. Another definition for mal-rules concept has been given at the start of this chapter. The amount and the type of mal-rules that students show in their works are the indicator or measure of their knowledge of the domain; because Greenhow reports the mal-rules might inform us the level of the facility or understanding that students display. Payne and Squibb, in their study, which has been mentioned above, reported that some mal-rules were "severely skewed "(p. 1). They also found that some mal-rules being irregular and some infrequent. These instability and variations in the errors, which students make, show the difficulty of classifying mal-rules. Similar conclusions have been made by many scholars. Walker et al. (2015) conducted case studies to build mal-rules using students' incorrect answers. In one of the cases, involving 110 students and 329 submissions, the DEWIS system was used to assess the students' answers. They compared the students' incorrect answers with the correct answers and determined the likely cause of errors and built mal-rules. They have also reported the impossibilities of describing some errors that had been uncovered. When interpreting some errors, they used conjectures. The challenges of identifying and classifying errors and/or mal-rules are analogous to that of solving complex problems.

Hardin (2003), Duncker and Lees (1945) and Mayer (as cited in Funke, 2010) describe the concept of a problem as an entity having three components, which are the current state(s), the desired state(s) and the process(es) or action(s) that are required to overcome the obstacle(s) that prevent(s) achieving the desired state(s) or outcome(s). Problems are categorised into ill-defined and well-defined (Hardin, 2003) groups; the concern of this review is about ill-defined, which is also called a wicked or complex problem. According to Hardin (2003) what makes a problem ill-defined (complex) is its features that involve the complexity in its portrayal and the needs for multiple means to solve it. A wicked problem also displays similar features as that of an ill-defined problem's characteristics. Most scholars refer to Rittel and Webber's (1973) identification when they discuss about a wicked problem. Rittel and Webber claim there are at least ten different reasons that can be ascribed to identify the complexity of a wicked problem. Rittel and Webber's recognition of a wicked problem's characteristics leads to a simplified definition, which states that a wicked problem has no clear mission and method(s) to reveal the purpose of the mission. Dörner and Funke (2017), who are the leading authors in the CPS research field, state that "complex problem solving is a collection of self-regulated psychological processes and activities necessary in dynamic environments to achieve ill-defined goals that cannot be reached by routine actions" (p. 6). Complex problems show varieties of composite characteristics. According to Dörner and Funke (2017), in a complex system (1) the abstraction is hierarchical, (2) the system is dynamic and develops overtime (3) and contains large amount of knowledge that is interrelated "together with a broad list of potential strategies (domain- specific as well as domain-general)"(p. 6). In the definition of a CPS, Yeo and Marquardt (2012), similarly, refer to the attributes in the CPS as being "highly unpredictable" (p. 263). Furthermore, Wüstenberg, Greiff, and Funke (2012) identify the CPS tasks displaying characteristics such as availability of insufficient information at the start of solving the problem, requirement from problem solvers to know suitable approaches to create information and the ability to apply procedural knowledge to be in charge of the problem or the system.

The CPS challenges exist in all fields. In sections 2.4.2 and 2.4.3 the limitations of CAA systems assessing subjective assessment has been identified. This limitation is recognised in the statement that says, "trying to make a modern computer do what it is supposed to can turn out to be a complex problem" (Fischer, Greiff, & Funke, 2012, p. 36). The difficulties of some mal-rules' identification and classification can also be categorised under the complex problem-solving topic or field; more detail about mal-rules creation complexities is given at the start of this section. Approaches to understand and/or solve a complex problem will be discussed next.

Sonnleitner, Keller, Martin, and Brunner (2013) and Fischer et al. (2012) stress that the problem solvers should acquire knowledge about the problem. Funke (as cited in

Wüstenberg et al., 2012) believes observing problems and solutions characters and obtaining knowledge to achieve a target are essential skills required in a problem-solving process. Yeo and Marquardt (2012) report that rule identification, which is “referring to the quality of the applied exploration strategy”, rule knowledge and rule application constitute the structure of CPS. Wüstenberg et al. (2012) explain that rule knowledge is the knowledge generated while the knowledge application is the ability to achieve the purpose. Hence, understanding the structure of CPS is helpful in comprehending the problem and finding partial or acceptable solutions. The DSR field also works in solving wicked problem in the area of artifacts development. According to (Hevner, March, Park, & Ram, 2004), “design-science research in IS addresses what are considered to be ‘wicked problems’ (Brooks 1987, 1996; Rittel and Webber 1984)” (p. 81). As such Hevner et al. (2004), list the characteristics of wicked problem in the DSR in IS as follows:

- Unstable requirements and constraints based upon ill-defined environmental contexts.
- Complex interactions among subcomponents of the problem and its solution.
- Inherent flexibility to change design processes as well as design artefacts (i.e., malleable processes and artefacts).
- A critical dependence upon human cognitive abilities (e.g., creativity) to produce effective solutions.
- A critical dependence upon human social abilities (e.g., teamwork) to produce effective solutions. (p. 81).

The wicked problems that happen in IS and CAA systems development research are similar because the purposes of the two research are producing information technology artefacts; therefore, CAA systems researchers can utilise the DSR approaches to conduct their research.

2.7 Discussion

The findings from the literature review are shown in Table 2.1. Each finding will be referred to by its finding number and discussed next.

Table 2.1 Literature review outcomes/findings

Findings Number	**Review's findings or outcomes**
1	Importance of academic assessment in students' learning life
2	Questions types impacts on the assessment characteristics
3	Assessment criteria controlling assessment qualities
4	E-assessment ability to assess objective questions
5	E-assessment limitation in assessing subjective questions
6	General e-assessment of mathematics system development methodologies
7	Relationships between e-assessment limitation, mal-rules and CPS
8	DSR dealing with CPS in IS context
9	Relationships between DSR artefacts and e-assessment of mathematics

Finding 1 is about the role of academic assessment in students' learning life, which determines the outcomes of the types of the knowledge they receive. It is revealed that "assessment should be a way of life – part of the institutional culture" (Keeling & Dungy, 2004, p. 26). Students learning can be influenced by the type of assessments that they are engaged with. For example, what formative assessment provides is different than what is achieved in summative assessment. Students learning can be improved by applying formative assessment, which influences the outcome of summative assessment (Kulasegaram & Rangachari, 2018). This discussion about the assessment roles in students' activities shows the significance of the academic assessment in knowledge production. Section 2.3.1 gives more details of the purposes of formative and summative types of assessment. To produce genuine knowledge the assessment characteristics are shaped by the types of questions they use.

Findings 2 is related to the objective and subjective questions revealing different levels of knowledge producing deep and surface learning. The comparisons between the two approaches are displayed in the students' aims and attitudes. In a study it was found, in deep learning approach a student "aim in the course was to develop their own understanding of the subject matter and the students reported being intrinsically motivated" (Asikainen, 2014, p. 40). On the other hand, in the surface approach, it was reported students "having little or no interest in the course and their sole aim in the course was to pass it with minimum requirements" (Asikainen, 2014, p. 43). The differences in the attitudes and aims that were directed by the two approaches show significant impacts on the life of the two groups. This example is another clear indication of the impact of assessment on students' achievement. In addition to the approaches, the assessment criteria also improve the assessment effective outcomes. The role of assessment criteria in controlling the assessment quality has been shown in finding 3. The implementation of concepts such as method marking, partial credit and follow through errors assist in revealing the examined domain knowledge adequately. Section 2.4.1 gives more details about what types of knowledge that method marking implementation can assess. The first three findings are parts of the facts about authentic assessment that is expected to produce genuine knowledge.

Findings 4 and 5 relate to the competency of e-assessment of mathematics assessing objective and subjective questions. The review and discussion of the limitation of e-assessment or CAA in assessing subjective assessment is the core of this literature review. This CAA limitation issue is concluded by well-known author and the developer of STACK CAA system; Sangwin (2015b) says "what we cannot do is easily encode the meaning of an expression and combine this with simple logic and automatic CAS calculations. Until we can achieve this simple interface, marking student's extended work automatically will be impossible" (p. 711). Sangwin's statement clearly identifies this issue as critically important requiring continuous search for the solution. Perhaps this research outcome might add some knowledge that can be contributed towards general research efforts that in future resolve this issue.

Finding 6 is about the methodologies that are used to develop CAA of mathematics system. The researcher acknowledges the impossibility of investigating each existing CAA systems. Therefore, the researcher has chosen to describe CAA systems generally in terms of the components these systems rely upon. The details of the development of these components have been shown in Figure 2.1 and Figure 2.2, and explained in section 2.5. One of the main concerns of the current research is to uncover the challenges that CAA systems developers face when developing these systems to assess subjective questions.

Finding 7 links these challenges to the complexity of CAA systems assessing subjective questions. These challenges are similar to a complex problem solving (CPS). The natures of such CPS and some mal-rules identification and classification have been identified in section 2.6. Knowledge accumulation about complex problems and their solutions as well as the ability of applying the acquired knowledge are the key factors to produce acceptable full, if possible, or partial outcomes.

Findings 8 and 9 are about DSR in IS research. The natures of Complex problems in CPS research and wicked problems in DSR have been discussed in section 2.6. The purpose of DSR paradigm, in IS research, is to conduct research about wicked problems to create DSR artefacts that are used in solving these wicked problems. Hence, the DSR paradigm is considered the most appropriate to conduct this research.

Finally, the findings that have been shown in Table 2.1 and the following discussion demonstrate achieving the purposes of this literature review that have been set at the outset of this chapter. It has been stated in section 2.1 that the outcomes of the secondary research determine or guide the primary research activities. Before explaining the activities in the primary research, it is important to explain rationales for not forming research questions or research problems based on the gaps that have been uncovered in this literature review.

In this research, section 1.1 has discussed the nature of this research problem and types of class of problems. It has been identified that this research problem is associated with the research purpose of the improvement quadrant identified by Gregor and Hevner (2013). Any DSR project that is associated with the improvement quadrant focuses on producing knowledge that is used in improving existing artefact or producing a new artefact. In both cases, the research problem is known in advance. Therefore, there is no need to form specific research questions or problems based on the gaps that have been unexplored in this literature review. In this research, the focuses are on investigating the complexity of assessing subjective questions when using CAA of mathematics systems through studying a relevant context and providing knowledge that assists in improving these systems in the form creating new artefacts.

The purposes of this research review are to understand the problem further and suggest means that assist in finding solutions. Numerous topics have been included in this review. For example, the complexity of assessing subjective questions using e-assessment tools is one of these topics. The complexity of this issue has also been discussed in chapter 1 and is demonstrated in chapter 5 when presenting and discussing the analyses of student workings. Furthermore, the findings from this review have been listed and discussed in section 2.7. The suggestions from the outcome of the discussion will be considered in the primary research part of this study.

The matters, which will be studied, in the primary research include the investigation of the context that resembles CPS to understand challenges of developing CAA artefacts. The context that is selected for investigation is student scripts. The student workings will be analysed to reveal solution strategies that student use and errors that are made during the application of these strategies; these strategies and errors will be examined and categorised. The significance of uncovering new errors in improving the utilities of CAA systems is stated by Sangwin (2013); he emphasises “identifying these, through professional experience, or collaboration with educational researchers, remains a challenge” (p. 5). He is referring to the specific misconceptions and technical errors, in student workings, in being instrumental in producing a particular answer in their workings.

In this research the main contributions of the research include the methodology to investigate the context, the creation of questionnaire as the research instrument and design knowledge in DSR artefacts.

Chapter 3 Research process

The findings from the literature review in chapter two have shown that assessing subjective questions using CAA systems is like a CPS. Moreover, the review has found that CAA systems are specific artefacts that could be considered as one of DSR outputs. The importance of applying DSR paradigm to solve CPS in IS research has also been acknowledged in the outcome of the review. The significance of searching for the knowledge that assists in developing and/or improving CAA systems has also been found. Therefore, a new context, which is student scripts, will be studied to reveal this knowledge.

Additionally, the importance of this knowledge to develop the DSR artefacts has been emphasized. Hence, establishing the relationships between producing and applying this knowledge, and design science, is significant.

This chapter intends to investigate the relationships between natural, social and design sciences research to identify the knowledge that is required to create DSR artefacts. Furthermore, the chapter focuses on examining some leading DSR frameworks to construct the specific DSR framework to be used in this research. The research philosophies (ontologies and epistemologies) regarding student scripts and DSR artefacts likewise will be investigated to choose appropriate assumptions, concerning DSR and scripts entities. The assumed philosophical stances will be applied to select the relevant methodologies that guide the selection of research methods to analyse student scripts and to construct DSR artefacts. The chapter is subdivided into sections to discuss, explain and implement the aims that have been highlighted above. The structure of the chapter and overviews of the sections are entirely discursive; all the references to the publications will be given in the subsequent sections.

Section 3.1 discusses the relationships between the three scientific fields. This section after identifying design science, explains the DSR needs for the knowledge that both natural and social sciences research produce. The roles of the natural and/or social science methodologies in the analyses of student scripts will be discussed. Additionally, the section describes the DSR need for the application of design science methodology to construct and evaluate DSR artefacts.

The researcher has selected five well known frameworks in IS research when DSR is used as the research paradigm. The investigation of these frameworks is detailed in section 3.2. Different phases or activities in the frameworks are grouped and discussed in relation to this research. At the end of the investigation, a DSR framework for this research will be formed.

The second phase in the DSR framework consists of two parts. The first part is about the research philosophies and related matters that are discussed in section 3.3. In general, the research philosophies are implicitly or explicitly present in scientific research. In section 3.3.1 different paradigms matters are discussed. The term ‘enquiry paradigm’ is identified in the context of scientific research and multiple definitions are attached to it. The existence of different versions of these paradigms are also discussed and the relevant version for the current research is determined. The importance of researchers’ declaration of their research philosophical stances is also mentioned in the section.

Scientific enquiry paradigms have components which are the philosophies of ontology, epistemology, methodology, axiology and methods. In section 3.3.2 the ontology of being is identified and the detailed discussion of the relevant ontologies that could be used to analyse the scripts are performed. The discussion resulted in identifying and selecting the ontologies that recognise scripts as the real-world objects. The definition of epistemology and the relationships between epistemology and ontology are presented in section 3.3.3. The purpose of epistemology, in particular knowledge creation, using an appropriate methodology, is also discussed; furthermore, the link between knowledge and the type of methods to create knowledge and the philosophical assumptions that have been adopted by the researcher are also discussed and established. The mechanisms to ensure the validity and reliability of created knowledge is also explained. At the end of these discussions, the 'critical realism' research philosophy is adopted to investigate scripts.

The issues of ontology and epistemology regarding DSR artefacts are discussed in section 3.3.4. In this section, after identifying the positions of paradigms concerning DSR and IS research, the DSR artefacts' ontological and epistemological stances are determined as being the real-world objects. The relationships between researchers and the artefacts are also identified as varying between 'objectivism' and 'subjectivism' positions. Based on the investigation, the appropriate research philosophy to construct and evaluate DSR artefacts will be determined.

Section 3.3.5 is about the research methodology which consists of general procedures that contain different types of methods. The explanations of the three well-known methodologies, which are 'qualitative', 'quantitative' and 'mixed methods', are given in sections 3.3.5.1, 3.3.5.2 and 3.3.5.3 respectively. In this section the differences between methodology and research process is also stated. The definition of axiology and its philosophical influence on the entire research is described in section 3.3.6. The axiology philosophy emphasises researchers' responsibilities towards their activities regarding the ethical issues such as value-free and value-laden positions. The process that is followed to obtain access to the scripts is likewise explained in the section. The explanation of the research philosophies is completed with the description of research methods in section 3.3.7.

The section also explains the relationships between methodologies and methods. In this research the two methods that have been selected, to analyse scripts, were content analysis and survey methods. The explanations and implementations of the two methods are detailed in chapters four and five. As for the design science methods, their explanations and implementation are covered in chapter six.

3.1 Natural, Social and Artificial Sciences Relationships.

This section examines the relationships between artificial, natural and social sciences in terms of the artificial science research utilisation of the knowledge produced by the natural and social sciences research. Vaishnavi and Kuechler (2015) identify the artificial science as "a body of knowledge about the design of artificial (man-made) objects and phenomena—artefacts—designed to meet certain desired goals" (pp. 52-53). Pries-Heje, Baskerville, & Venable (as cited in Shedlock, Vos, and Chard, 2016) and Beckman et al. (as cited in El Idrissi and Corbett, 2016) also point to the same purpose of artificial science research defined by V. K. Vaishnavi and Kuechler. Artificial science is also called design science (Dahlbom, Beckman, & Nilsson, 2002). Different types of artefacts are produced for different goals. One of the disciplines that is closely related to design science research (DSR) is 'information systems' (IS) research. There are multiple definitions associated with the 'information systems' term. However, here only one defection, which is relevant for this research project, is given. According to Alter (2008), IS "is a system in which human participants and/or machines perform work (processes and activities) using information, technology, and other resources to produce informational products and/or services for internal or external customers" (p. 6). Lyytinen, Newman and Mumford (as cited in Haki, & Aier, 2015) report that information technology (IT) artefacts are one of the components that constitute IS.

In DSR to create artefacts, the presence of other sciences' knowledge is necessary. For example, Shedlock et al. (2016) use the example of culture having influence on the design of artefacts and vice versa. The knowledge that comes from "social or behavioural (human) science"(Gregor, 2010, p. 58) helps in understanding the characteristics of communication between artefacts and their users (Gregor, 2010); this knowledge, as Hevner et al. (2004) view, “should focus on truth” (p. 82) which “informs design” (p. 82). Gregor also has the view that knowledge produced by behavioural science research assists in designing the artefacts. Furthermore, the IT artefacts creation obeys the “physical laws” (Gregor, 2010, p. 53) that, as Gregor, explains emanate from the knowledge in the natural science. On the other hand, the relationship of other sciences fields to the artefacts are also important. In this regard, Beckman, Nilsson, and Dahlbom (2002) state that the artificial science is “a concern to all branches of science” (pp. 52-53) thus emphasising the association between artificial science and other sciences. March and Smith (1995) also identified the scientific interest in information technology in two areas; explaining IT phenomena using descriptive research to understand the nature of IT; and prescriptive research to improve IT performance. The aim of descriptive research, which belongs to natural science, is to produce knowledge while prescriptive research, which is originated in design science, is to use knowledge to produce artefacts. The above discussion shows the dependency of the DSR on the knowledge produced by the other sciences; additionally, the investigation has revealed the relationships of other sciences in the knowledge that is deduced from the DSR artefacts investigation and/or creation.

Scientific methodologies are required to produce scientific knowledge. Natural and social sciences research use only behavioural and/or natural research methodologies. However, DSR requires both behavioural and/or natural and design science methodologies to create artefacts. Ken, Tuure, Marcus, and Samir (2007) state that natural and social sciences methodologies are “not often applicable to the solution of problems encountered in research and practice” (Ken et al., 2007, p. 47); the issues these authors are talking about are IS research and practice that involve building real artefacts. And for IS research to be relevant, its research output should contain explicitly applicable solutions that resolve the IS communities’ problems (Ken et al., 2007; Venable, 2006). Therefore, the lack of natural science research methodologies to realise the aim of DSR leads to the usage of both natural and design methodologies. Further details that show the relationships between design, design science and DSR topics will be given in section 6.2.

DSR scholars recommend reviewing some of known DSR methodologies or frameworks; because the familiarity with the current DSR methodologies helps “researchers to present research with reference to a commonly understood framework” (Ken et al., 2007, p. 48). Thus, section 3.2 focuses on investigating the selected DSR frameworks and constructing an appropriate framework for this research.

3.2 DSR Structures

This section explains the purpose of a DSR structure and means to develop one for this research by reviewing the selected DSR structures. The purpose of a ‘research structure’ is to enable researchers to achieve the research aim(s) in a systematic way; such a system will be a research process. Nunamaker, Chen, and Purdin (1990) state “a research process involves understanding the research domain, asking meaningful research questions, and applying valid research methodologies to address these questions” (p. 91). Different research fields develop their own research structures or adopt one from related discipline(s). Information systems research, which uses the DSR paradigm, adopts the methodologies from engineering. There are several existing structures (frameworks) to conduct IS research using the DSR paradigm. In this section, five different well-known structures, (which have been published in five different papers), will be discussed. In this section the research structure phrase will be used when discussing these structures. The research structure phrase refers to the ‘research framework’.

The research structures in the five selected papers are made up of components or activities or phases. Some papers present the phases in the sequential order that the DSR process requires to be carried out to realise the research intention, and some other papers show steps to be taken without specifying the order of these steps’ application; in this case, researchers are required to determine their appropriate orders to help them achieve their research aims.

Four of the reviewed papers, (or structures), display the activities or steps in a sequential order; Nunamaker et al. (1990), Ken et al. (2007), Vaishnavi and Kuechler (2015) define the DSR research structure in a sequential order; Walls et al.'s (1992) also shows sequential order when creating information systems design theory (ISDT). However, Hevner et al.'s (2004) seven guidelines are not in the sequential order to guide IS research when DSR 'research structure' is considered. Regardless of the order of the steps/phases in the structures that are presented in the reviewed papers, the relevant subjects have been searched and evaluated; the knowledge inferred from the results of evaluation has been used to construct the framework for this research.

3.2.1 Phase 1: Research problem identification and relevancy.

In the selected publications or papers, this section generally, considers (1) the actions to be taken to identify research problems, (2) the justification of the research outcome and (3) the identification of research problems sources.

The first step in the research structures of Nunamaker et al. (1990), Ken et al. (2007) and Vaishnavi and Kuechler (2015) is to examine research problem(s); similarly, Hevner et al.'s (2004) second guideline emphasises this step which is called "problem relevance" (p. 84). Nunamaker et al.'s first phase and Hevner et al.'s second guideline present further details than Ken et al. and Vaishnavi and Kuechler's problems identification phase; on the other hand, Walls et al. (1992) address this issue with less detail. Nunamaker et al. and Hevner et al. state the usage of natural or social science to study a phenomenon to understand research problems further; and to find solutions and produce knowledge that can be used in developing artefact(s). Nunamaker et al. suggest the creation of the conceptual framework that might include the design that requires the development of artefacts to test the validity of the design. In the second guideline, Hevner et al. expect the knowledge to include, in addition to the identification of research problems, the definition of goal criteria and constraints on the artefacts to be developed. The knowledge produced at this stage, according to Nunamaker et al., could be used to build theory that contributes to the artefacts' development. However, Hevner et al. do not state the activities to output a theory.

Likewise, Ken et al. (2007) and Vaishnavi and Kuechler's (2015) DSR structures start with the research problems' determination; they specify this stage as the "problem identification and motivation" (Ken et al., 2007, p. 46) and the "awareness of problem" (Vaishnavi, Kuechler, & Petter, 2017, p. 27) respectively. The purpose of this activity is to identify a "specific research problem and justify the value of a solution" (Ken et al., 2007, p. 12); and the solution Ken et al. refer to is the artefact. This is similar to Hevner et al.'s (2004) second guideline and Nunamaker et al.'s (1990) initial system development stage. Ken et al. claim the knowledge required in this step is the "state of the problem and the importance of its solution" (Ken et al., 2007, p. 55). However, Ken et al. do not explain how to create this knowledge. Similarly, Vaishnavi, Kuechler, and Petter's (2017) first step is identifying the research problems from sources such as "new developments in industry or identification of problems within a reference discipline" (Vaishnavi, Kuechler & Petter, 2017, p. 8) and findings of the relevant disciplines to form proposals. According to Vaishnavi, Kuechler, and Petter, the research proposal should include "a Tentative Design and the performance of a prototype based on that design" (Vaishnavi, Kuechler & Petter, 2017, p. 9); the design should be a representative of the problem solution or at least the proposal should be used as the starting point to find solutions for the research problem(s).

On the other hand, Walls et al. (1992) ISDT building starting point is different. The ISDT building stages are depicted in Figure 3.1. In this figure, there are two parts that the term design is concerned with; these are the design process and the design product. The first component in the two parts is called ‘kernel theory’ which comes from the “natural or social sciences which govern design requirements” Walls et al. (1992, p. 42).

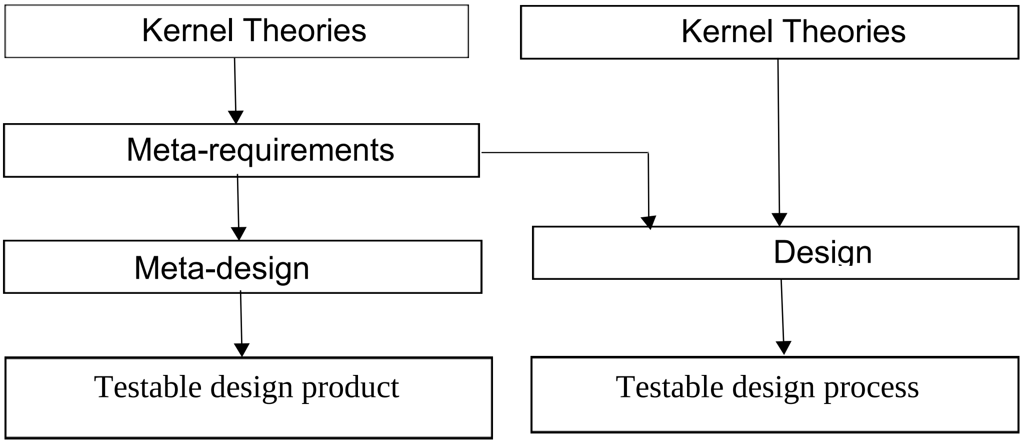


Figure 3.1 Relationships among ISDT components - Reproduced from Walls, Widmeyer, and El Sawy (1992, p. 44)

This shows, implicitly, the necessity of building theories, if not in existence, as part of creating ISDT. sciences which govern design requirements” Walls et al. (1992, p. 42).

The natural or social studies that produces theories must be based on the relevant research problem(s); this issue is also not stated in Walls et al. but can be understood from the general academic research tradition’s prerequisites that entail the examination of research questions and the existence of interest in the research to be conducted.

From five papers that have been reviewed, four papers considered the ‘problem identification and problem relevancy’ as the first activity in IS research using DSR paradigm. Walls et al. (1992) ISDT starting stage, as explained above, can also be traced back to the initial stage shown in the other three papers. The examination of the first phase in four papers and the second guideline of Hevner et al.’s (2004) has shown the difficulty of extracting similar knowledge from these papers. These differences have been found in all phases that have been examined. However, the important outcome from the review of these papers is extracting the relevant knowledge that is useful for this research.

The first step taken in this research was forming and presenting the research proposal. The proposal was produced using the literature review method, i.e. the outcome of the literature review was the main source used to identify the research problems and to determine the relevancy of the problem to be DSR ones. When compared to the reviewed papers, the current research's first step is partially similar to Vaishnavi and Kuechler (2015) and Ken et al.'s (2007) initial stages in producing the research proposal.

3.2.2 Phases 2 and 3: phenomenon investigation and conceptualization.

The discussion about the second and the third phases is based on Ken et al.'s (2007) second activity which states "defining the objectives for a solution"(p. 55), Vaishnavi and Kuechler's (2015) "suggestion", Nunamaker, Chen, and Purdin's (1990) "develop a system architecture" (p. 99) and Walls et al.'s "meta-requirement" identification. Ken et al. suggest taking actions that involve using the identified problems and "knowledge of what is possible and feasible" (p. 55) to deduce solutions. The knowledge required at this stage, besides the knowledge obtained in phase one, is existing solutions' knowledge. Vaishnavi and Kuechler's suggestion phase is similar to the one that is identified by Ken et al.; Vaishnavi and Kuechler suggest using existing or new and existing knowledge to add new functionality to that envisioned in the original tentative design or solution. The intentions of a solution can be achieved through purposeful functionalities of an object or objects. In this phase, Nunamaker et al. activities are divided into two parts; the part that is related to the systems' (artefacts') components identification, the determination of relationships between these components, and the part that is concerned with the development activities, such as identification of artefacts' functionalities, constraints on the development of the artefacts, and objectives of the artefact's development; this part also includes developing evaluation criteria, designing and implementing the artefact. Hevner et al.'s second guideline covers the identification of research goals and constraints on the artefacts to be developed. The second phase in Walls et al.'s (1992) ISDT is to drive meta-requirements from the kernel theory; the activities in this phase is similar to the activities in the four papers. At this stage, the activities in Nunamaker et al. include actions that are made in the development phase in contrast to the other papers proposed actions. It is also noticeable that Ken et al.'s second activity is one step behind the activities identified in other papers. Ken et al.'s purpose is to identify the goals of the research not the functionalities of the artefacts.

This research second phase implements part of Nunamaker et al.'s (1990) first phase, Vaishnavi and Kuechler's (2015) suggestion phase, Ken et al.'s (2007) second activity and Hevner et al.'s (2004) second guideline. In this research, following the approval of the research proposal, a relevant context to understand the research problem was analysed, and the outcome of the analyses was used to produce foundational knowledge

3.2.3 Phase 3 DSR Artefacts Design

In this phase all the papers focus is on establishing functionalities, generating different design alternatives, implementing the chosen designs and preparing data to be used in the artefacts' creation.

Ken et al.'s (2007) purpose, as the third activity, is establishing functionalities, generating and implementing designs to create artefacts. Ken et al. state the presence of knowledge in the form of theory to perform activities in this phase. Vaishnavi and Kuechler's (2015) development phase focuses on enhancing the tentative design to construct artefacts. However, Vaishnavi and Kuechler claim that the knowledge, at this stage, does not have to reach the level of theory to be suitable for creating artefacts. The knowledge being in the form of theory is what Nunamaker et al. (1990) suggested in phase one. In this phase, Nunamaker et al. (1990) propose the determination of the design of data structure and database or knowledge base elements, forming various design alternatives, exploring and evaluating the design alternatives, determining and choosing the better design solution from the alternatives presented and specifying “the program modules and functions” (Nunamaker et al., 1990, p. 100).

Hevner, March, Park, and Ram’s (2004) sixth guideline which is “design as a search process” (Hevner et al., 2004, p. 88) focuses on designing artefacts. The main activity in this guideline is the build and test cycle. Artefacts are built using available means by taking into consideration constraints called ‘laws’; alternative design solutions are used to overcome or satisfy existing and new uncovered laws. Using all or any available resources, goals (called ‘end’ by Hevner et al.) and laws are transformed to abstraction and representation. The build and test (evaluate) cycle is applied to refine the design artefact to produce an acceptable research result. The explanation of the functionalities of these artefacts and the identification of the environment in which they are used in is one of the activities in this guideline. Walls et al. (1992) design method and meta-design use meta-requirements, which are obtained from kernel theories, to output a design product or design process.

The central issue in the discussion of these papers is to find the most relevant method to produce DSR artefacts which are regarded design products; this is achieved according to Walls et al. (1992) by employing a meta-design which “describes a class of artefacts hypothesised to meet meta-requirements” (p. 43). However, the concern, in this DSR research structure review, is about the methodology to create one type of artefact (not class of artefacts); although phases or steps in research structure to create a class of artefacts and one type of artefact might be similar. Thus, all the papers contents in the third phase give detailed DSR development methodologies, although, Walls gives only a brief explanation.

The third phase in this research extensively reanalysed the knowledge produced in the second phase. The combination of the knowledge from the analyses of student workings and other relevant knowledge have been used to form possible functionalities to achieve the research goals; the main guidance in forming functionalities was kernel theories while the knowledge inferred from the analyses of student workings was the main controller or the determinant of what is possible to apply in achieving the listed functionalities; multiple design solutions were explored to produce the knowledge that contributes to understanding the problem and assist in creating the DSR artefacts.

3.2.4 Phase 4 Evaluation.

The evaluation process, also called testing, contains activities that require multiple steps using appropriate method(s). Hevner et al. (2004), Vaishnavi and Kuechler (2015) and Ken et al. (2007) reveal that evaluation activity starts with preparing metrics that are based on the knowledge used in constructing the artefacts; Hevner et al. identify multiple methods that are used to employ these metrics; the refining of metrics, their application to test the artefacts and feedback from evaluation continue till the satisfactory research aim(s) have been achieved. In Vaishnavi and Kuechler's DSR process model, the evaluation criteria were initiated in the problem awareness phase; in the third guideline, Hevner et al. specify the source of metrics as being the requirements that are provided by the environment and the knowledge that resides in the research domain. Ken et al.'s (2007) fourth activity, termed "demonstration" (Ken et al., 2007, p. 55), performs limited evaluation that encompasses the artefact solving "one or more instances of the problem" (Ken et al., 2007, p. 55); in fifth activity Ken et al. performs full evaluation which is similar to the ones explained in other papers. However, Ken et al. do not state explicitly the sources of the knowledge for developing the metrics. Hevner et al. and Nunamaker et al. propose various types of evaluation methods. Walls et al. (1992) evaluation method is based on testing hypotheses that are developed as part of evaluating ISDT in which the developed meta-design achieves meta-requirements. Thus, the cited papers comprise comparable evaluation activities, except Walls et al.'s paper, with the sufficient details that can be adopted in the evaluation of this DSR artefacts.

In this research, qualitative, observation and testing evaluation methods have been used to evaluate DSR artefacts; furthermore, there are other evaluation methods that have been used at different stages of this research.

3.2.5 Research contribution.

There is no specific phase that is associated with the research contribution. Generally, the expected contribution is assumed at the start of research, but the actual contribution is assessed at the end of the research. However, the activities that are associated with knowledge acquisition covers the entire research duration. The process to reveal the contributed knowledge involves applying relevant knowledge acquisition methods to uncover the knowledge to be contributed. Furthermore, this process- revealing contributed knowledge- requires specifying and applying validation criteria to verify the claimed knowledge. The knowledge acquisition methods are selected from existing research methodologies- qualitative and quantitative (mixed methods) - that were applied to analyse student scripts and DSR design science method to produce DSR knowledge. Vaishnavi and Kuechler (2015) define abduction, deduction, reflection and abstraction methods to acquire knowledge at different phases of DSR research structure using the knowledge cognitive process. More information about the cognitive process is given in Vaishnavi, Kuechler, and Petter (2017, p. 10). In contrast, Hevner et al. (2004), Nunamaker et al. (1990) and Ken et al. (2007) do not provide a specific method to acquire the contributed knowledge, but it can be implicitly understood from their views of behavioural and design science methodologies activities. Nunamaker et al. indicate the developed artefact functions “both as a proof-of-concept for the fundamental research and provides an artefact that becomes the focus of expanded and continuing research. Contributions at each stage of the life cycle obviously contribute to “fuller scientific knowledge of the subject”” (Nunamaker et al., 1990, p. 92). Nunamaker et al. do not point to specific activities within one cycle, for example, where during the construction of the artefact, contributions can be identified; in contrast, Ken et al. state that the “research contribution is embedded in the design” (Ken et al., 2007, p. 11). In the fourth guideline, Hevner et al. explain the sources of research contribution as being design artefact, design construction and design evaluation knowledge parts. Hevner et al. specify artefacts, foundations and methodologies as the main DSR knowledge contributions; they state the “representational fidelity and implementability” (p. 87) as the criteria to assess the claimed contributions. According to Hevner et al., representational fidelity refers to the concept that, for example, a model featuring exactly the object or the concept it represents, while the implementability refers to realizing the utility that the artefact is built for. The two criteria are demonstrated in building and evaluating the artefacts. The instantiation of the artefacts and demonstrating their utilities through evaluation is the proof of the claims made. In summary, the information that comes from the knowledge base and the researchers’ judgement are the sources that assist in (1) revealing the portions or activities that contain the contributed knowledge, (2) forming the knowledge validation criteria and (3) selecting appropriate knowledge acquisition methods or creating new methods to extract knowledge.

3.2.6 Phase 6 communication.

The absence of agreed upon standardized research presentation structure using DSR paradigm, in IS research, leads some researchers to present their works using other disciplines’ structures. In the examined papers, Nunamaker et al. (1990) and Walls et al. (1992) do not cover the subject of communicating the outcome of research. Vaishnavi and Kuechler (2015), Hevner et al. (2004) and Ken et al. (2007) discuss this issue with varying details.

Regarding the presentation of design research works, when compared with other research fields, “the design community lags the behavioural community in having useful templates for communicating knowledge contributions” (Gregor & Hevner, 2013, p. 349).

Therefore, it is important for researchers to have the knowledge that determines the positioning and communicating of their research efforts in order to convey the contribution of knowledge and theory to design science field and communities (Gregor & Hevner, 2013; Vaishnavi et al., 2017).

The type of contributed knowledge should be well known to the communities who access the DSR outcome. These types of contributions as well as the knowledge acquisition methods have been explained in the knowledge contribution explanation phase. According to Hevner et al.'s (2004) guideline seven, the extent of the communication should allow the audiences to reconstruct and evaluate the artefacts. Such details include all the descriptions of all the activities that are covered in the research. The structure and the contents of the communicated subject matters are based on the intended audiences. Ken et al. (2007) suggest the communication to include "the artefact, its utility and novelty, the rigor of its design and its effectiveness to researchers and other relevant audiences such as practicing professionals" (p. 56). For scholarly research publication, they recommend using six steps which are: (1) problem identification and motivation, (2) defining the objectives of a solution, (3) design and development, (4) demonstration, (5) evaluation and (6) communication. Ken et al. propose the structure that follows this order applying the structure sequentially or start from any phase and move outward. Ken et al.'s view-point, to determine the starting stage, is dependent upon the nature of the approach. For example, a problem-centred approach follows the sequential order and applies all the six stages identified above. The problem-centred approach is applied when "the idea for the research resulted from observation of the problem or from suggested future research in a paper from a prior project" (p. 56). Vaishnavi, Kuechler, and Petter (2017) and Vaishnavi and Kuechler (2015) discuss in their conclusion phase the categorisation of the knowledge gained from research and the identification of contributed knowledge; the details of knowledge acquisition methods and the type of knowledge obtained should be clearly communicated to the relevant audiences (Vaishnavi, Kuechler, & Petter, 2017).

Hevner et al. (2004), Gregor and Hevner (2013), Ken et al. (2007) and Vaishnavi, Kuechler, and Petter (2017) have proposed different formats for presenting contents and structures of DSR for IS research works. Ken et al. do not see differences in the structure of a DSR publication when compared to the publication of other research fields, such as natural science. Gregor and Hevner acknowledge the variances between the DSR and other research field's research presentation structures; therefore, Gregor and Hevner propose a different structure; Vaishnavi, Kuechler, and Petter (2017) do not detail their own structure but refer to Gregor and Hevner to articulate the significance of the research presentation issue.

Given the lack of a standard structure for communication research of this type, the structure of the DSR framework as shown in Fig 3.2 is use for guiding the communication this research project. Furthermore, the structure of this communication has adapted some of the suggestions that were made by Gregor and Hevner (2013) regarding the communication of DSR project.

3.2.7 Research rigor.

Research rigor is not a phase or part of a phase. It is applied to all activities in the research to assist in producing a valid research output. It involves searching, selecting and applying appropriate extant knowledge from the knowledge base in research. Only Hevner et al. (2004) cover the issue of the research rigor in detail; They address it in the fifth guideline. To ascertain the importance of this guideline, other publications besides Hevner et al. have also been included in this review.

In a design science research project, the application of rigor is not only important in designing artefacts, but also in all activities of the design research project (Costa, Soares, & de Sousa, 2016). For example, Higgs, 2001 and Rice & Izzy, 1999 (as cited in Fereday and

Muir-Cochrane, 2006) report that the explanation of theoretical rigor involves rigorous and extensive discussion and reasoning to choose the most appropriate methods. This is also the case at the methodological level; Horsfall, Byrne-Armstrong, and Higgs (as cited in Fereday and Muir-Cochrane, 2006) explain that rigor in qualitative research, requires thorough designing, cautious consideration of studied objects, and obtaining of creative and purposeful outcomes. Hevner (2007), in the three design science research cycles framework, explains that the purpose of the rigor cycle is to connect design science activities to the knowledge base. According to Drechsler and Hevner (2016), in artificial design, the employment of rigor ensures the utilisation of the knowledge base which includes "scientific theories … experience and expertise" (p. 3). The relationships between research activities and the knowledge base is bidirectional; because research does not only use the knowledge that is in the knowledge base, but also adds novel knowledge, which is the result of research efforts, to the knowledge base (Jannaber, Riehle, Delfmann, Thomas, & Becker, 2017).

In a DSR project, rigor practices requires the research outcome not to be considered just routinely constructed efforts, but the results of scientific research efforts (Iivari, 2007); Iivari also states for researchers to claim that the artefacts development process is design science activity, they have to show the transparency regarding the artefacts development process. Referring to validating theory in prescription-driven research, Aken (2004) proposes "to do some weeding … with the help of some rigorous testing and grounding" (p. 240). Similarly, since "design science offers prescriptions and creates artefacts that embody those prescriptions" (March & Smith, 1995, p. 254), these same rigorous testing and grounding also apply to DSR works, if credibility of these works are wanted. Thus, the arguments that have been shown above, regarding rigor, emphasise the significance of rigor activities at all phases of a research project.

The rigor was applied in all phases of this research. In (1) the examination of the research purpose and research questions to ensure that they reveal research characters, (2) in the literature review process to understand the research problems and build research foundations, (3) in the understanding of research philosophies and adopting the relevant ones to select and apply the appropriate methodologies and methods to analyse student workings and using the results of analyses as the foundation to create the prescribed artefacts and (4) in concluding the project by communicating it in the way that enables the researcher to show the knowledge contribution of the research to the intended communities.

In summary, the output of the review of the scholarly publications has revealed the importance of examining selected DSR frameworks before conducting IS research using DSR paradigm. This type of a review broadens the knowledge of the context that is DSR. This knowledge has contributed in creating DSR framework for this research.

3.2.8 Adopted DSR framework.

The structures of DSR in the reviewed publications show phases in DSR in IS research. For example, design activities are performed in the design phase. After the review of each related topics' contents, the knowledge that was obtained from the reviewed papers was explored, discussed and the decision was made whether to adopt the reviewed component's in the DSR structure or create an alternative component that satisfies this research needs; for example, the reviewed artefact's design method has been adopted as the current research's artefact design method; because it showed relevancy to this research artefact design needs. However, if the publications did not contain the knowledge that satisfies this research requirements, then the outcome of the review and applicable requirements have been used to form an alternative component for that phase in the framework. Finally, the created components have been used to build the DSR framework, which is shown in Figure 3.2, named DSR artefacts for method marking (DSRA-MM).

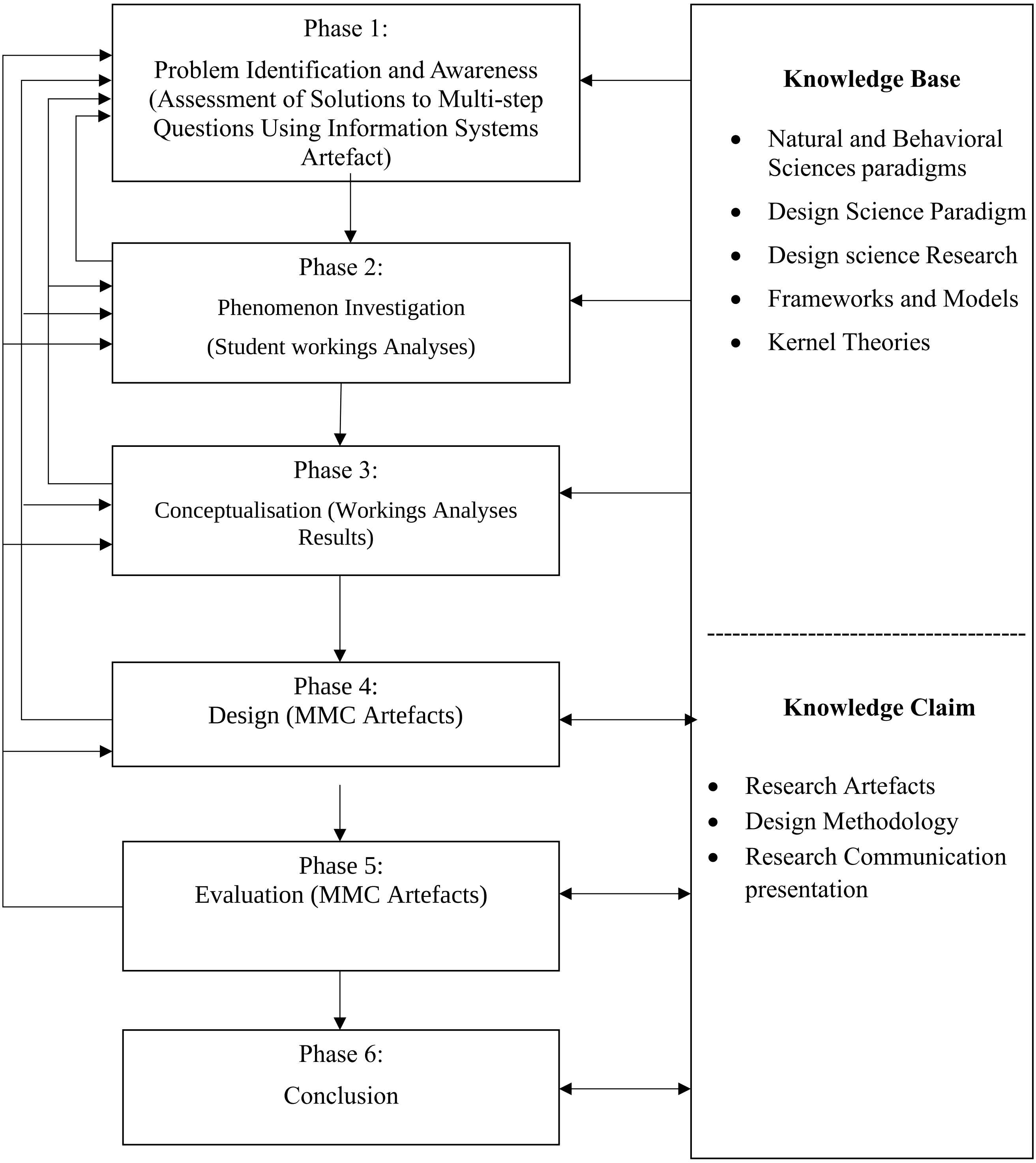


Figure 3.2 Method marking concept artefact design science research structure

The DSRA-MM framework is made up of six phases. The first phase is the research problem identification and relevancy phase that has been explained in section 3.2.1; the activities and tasks in this phase have been described and/or implemented in chapter two. The second phase is the phenomenon investigation, which has been explained in section 3.2.2. This phase consists of two main parts. The first part considers research philosophies, which will be disused in section 3.3, and its subsections. The second part focuses on the application of the assumed philosophies; and the details of this part will be covered in chapter four. The third phase is the conceptualisation of the findings, which is discussed in chapter five. The fourth phase is the design of the artefacts, which has been identified and explained in section 3.2.3. The tasks in this phase will be examined, implemented and described in chapter six. The fifth phase is the evaluation phase and the explanation of this phase has been given in section 3.2.4. The details of implementing this phase will be given in chapter seven. The sixth phase is the conclusion phase that includes the research knowledge contribution and the research communication. The descriptions of knowledge acquisition methodologies and research knowledge contributions have been covered in sections 3.2.5. The knowledge to be contributed will be identified and discussed in chapter eight. Finally, this research is communicated, using the approach that has been determined in section 3.2.6.

3.3 Research Philosophies for student workings and design science artefacts

The DSRA-MM framework that has been shown in Figure 3.2 presents the processes of analysing student scripts (workings) in phase two. In this phase, there are two main parts, which are the research philosophies and analysing student workings. The analyses of student workings will be discussed in chapter four. In this section, the discussion will focus on examining the existing research philosophies to select the relevant ontological and epistemological stances regarding student scripts to choose the relevant methodologies.

Furthermore, this part includes the examination of different stances of ontology and epistemology regarding DSR artefacts.

This section starts with the examination and explanation of the inquiry paradigms that are presented in section 3.3.1. The next discussion is about the ontology and this is covered in section 3.3.2. After the completion of the discussion of ontology, section 3.3.3 identifies and describes epistemology. The following section which is 3.3.4 covers the ontology and epistemology in DSR. The completion of the discussions about ontology and epistemology leads to the discussion of methodology which is given in section 3.3.5. The last two sections, 3.3.6 and 3.3.7, cover axiology and research methods respectively.

3.3.1 Research paradigm.

The following discussion is focused on the scientific or ‘inquiry paradigms’. In research, the paradigm’s concept is known with multiple terminologies such as “philosophical worldviews” (Creswell, 2013, p. 35) and “philosophical perspectives” (Moon & Blackman, 2014, p. 1173). The paradigm’s concept also has multiple definitions.

According to Guba and Lincoln (1994), “i*nquiry* paradigms define for *inquirers* what it is they are about, and what falls within and outside the limits of legitimate inquiry” (p. 108). Paradigms are also defined as assumptions, beliefs, values, rules, and guides. to help solve problems and/or perform actions (Guba, 1990), and ”in a more general sense it is the way we "see", "hear", and "touch" the world not in terms of the physical senses of sight, hearing”(Mink, 1992, p. 21). Other authors, for example, Kuhn (1970), Gorard (2010) and Hathaway (1995) limit these beliefs to within a particular scientific field. Gorard and Hathaway regard paradigms as the lenses, which are the collection of agreed upon rules belonging to a discipline, which researchers use to understand and solve single or multiple (as Gorard labels them) ‘puzzles’. Kuhn’s two views of paradigms are likewise mentioned here. According to Kuhn, paradigm

> stands for the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community. On the other, it denotes one sort of element in that constellation, the concrete puzzle-solutions which, employed as models or

> examples, can replace explicit rules as a basis for the solution of the remaining puzzles of normal science. (p. 175)

The two versions of Kuhn's (1970) paradigms that have been elaborated here fall within 'shared community' and similar 'puzzles' within the normal science; they do not cover all other disciplines to be regarded as worldview as other authors identified. The issues of variations in the paradigm's definitions are due to the paradigms' application scopes and matters involving the epistemological and ontological assumptions.

Regarding the paradigms' scope, Morgan (2007) groups them into four different related versions; these are (1) 'paradigms as model examples of research', (2) 'paradigms as shared beliefs among members of a specialty area' (3) 'paradigms as epistemological stances' and (4) 'paradigms as worldviews'. More details about these versions of paradigm are given in Morgan (2007, pp. 50-54). According to Morgan's classification of paradigm, DSR paradigm belongs to the 'paradigms as epistemological stances' type. The reason is that the DSR paradigm is specific to one particular field having one aim, which is producing artefacts. Therefore, it is regarded as a community related paradigm that has not yet reached the world view paradigm level.

The other reason that causes variations in paradigm definitions is due to the ontological and epistemological assumptions referring to both identification and usage in research. Blaikie (2007) relates the differences in paradigms to two "distinguishing characteristics that form their core on which they differ: namely… ontological … and … epistemological assumptions" (p. 13).

Many prominent authors also define scientific paradigms in terms of their philosophical components. For example, Creswell and Creswell (2013) and Greene and Hall (2010) define paradigms in terms of ontology, epistemology, methodology and axiology.

However, Guba (1990) and Blaikie define paradigms in terms of ontology, epistemology and methodology ignoring axiology. Silverman and Marvasti (2008) define paradigms as models that involve only ontology and epistemology. However, despite the variations in definitions, the consensus among these scholars is that paradigms contain components that are interrelated. Thus, researchers need to be familiar with different types of paradigms and their components to help them clarify their positions and/or perspectives they bring with them to the bodies of research. These definitions of paradigm show the importance of examining philosophies such as ontology and epistemology, which will be covered in the following sections.

3.3.2 Ontology.

The two types of beliefs about ontology are evolving ontology and permanent ontology; the discussion in this section is about the permanent ontology that is termed ‘Ontology of being’; as the phrase indicates, it is “the study of being” (Gray, 2004, p. 19), “the study of being” (Crotty, 1998, p. 10) with the main intention of revealing “the nature of reality” (Creswell, 2014, p. 54), “nature of reality” (Ritchie, Lewis, McNaughton Nicholls, & Ormston, 2014, p. 4), “nature of reality” (Guba & Lincoln, 1994, p. 108) and “nature of reality” (Saunders, Lewis, & Thornhill, 2009, p. 10) or ontology is “nature of what exists” (Gray, 2004, p. 19) and “what constitutes reality” (Gray, 2004, p. 19). Biesta (2010) identifies ontology as “the assumptions and beliefs we hold about reality and, more specifically, about the reality that is the object of research” (p. 11). The ontological assumptions enable researchers to distinguish research objects as independent or constructed reality or having both - independent and constructed- forms of existence (Crotty, 1998; Guba, 1990; Moon & Blackman, 2014; Ritchie et al., 2014).

There are varying degrees of ontological stances which are mainly divided into two opposing categories commonly called ‘relativism’ and ‘realism’ (Blaikie, 2007; Chen, Shek,

& Bu, 2011; Holden & Lynch, 2004). In this research, extensive evaluation of multiple types of ontologies have been made; however, only ‘shallow realist’ and ‘critical realism’ ontologies will be discussed in this section. The researcher believes both philosophies stances’ involvement in analysing student scripts assists in producing valid acceptable knowledge.

3.3.2.1 Shallow realism.

Shallow realism is the most extreme form of ontological realism that excludes human view in shaping the meaning of entities. Shallow realism is, as Bracken (2010) states, “the belief that the world of social interactions exists independently of what I perceive it to be, it is a rational, external entity and responsive to scientific and positivist modes of inquiry” (p. 2). ‘Positivist’ paradigm is associated with the ‘objectivist’ epistemology of which shallow realism ontology is its companion. According to Guba and Lincoln (1994) and Blaikie (2007), shallow realists believe that comprehension of facts about a phenomenon originates in investigating the phenomenon by applying an appropriate method (Moon & Blackman, 2014) in which the “researcher adopts objective procedures” (Blaikie, 2007, p. 60). In the views of objectivists,

> observations are expected to be neutral and represent no particular interest or purposes; descriptions, likewise, are to be an objective or detached report of what happened. To demonstrate one knows reality (or subject content) one must accurately describe it or reproduce it. A statement is true when it corresponds to reality as empirically validated, and false when it does not. (Pratt, 1998, p. 22)

The type of knowledge that is claimed by Pratt (1998) above, is not constrained by a context or time and the way it is perceived is governed by “immutable natural laws and mechanisms” (Guba, 1990, p. 109); and Blaikie (2007) clarifies this by claiming “what you see is what is there” (p. 14). The shallow realism ontology only allows the exploration of “what is, without examining what is behind what is” (Stickley, 2006, p. 573). The answer to the “what is” question does not require generating, for example, why an event takes place but requires the knowledge that shows the existence of the event itself. It does not involve scanning for causes and effects to answer why questions; however, this knowledge is necessary to explore things ‘behind what is’.

The importance of determining the existence of the ontology to be studied, or according to Harré and Madden (as cited in Roberts, 2014) and Tilly (2008), the act of neglecting the establishment of the presence of physical and social research realities before subjecting them to researchers’ observation and interpretation risks the produced knowledge being misleading and incorrect or a “false conclusion” (Tilly, 2008, pp. 5-6); this would lead to the presumption that the lack of a reliable foundation for the produced knowledge would result in rejecting the knowledge as true or acceptable knowledge. That is why according to Blaikie (2007), the shallow realist ontology is suitable to analyse research objects using an inductive research method. Further details about the inductive method is given in section 4.2. Hence taking the shallow realist ontology stance when analysing the content of student scripts using qualitative methodology is appropriate for uncovering themes relating to errors and solution strategies. Student workings are independent entities and applying inductive qualitative content analysis can reveal the mentioned knowledge without the researcher’s interference in the outcome of the analysis. There is no interpretation, at this stage of the analysis, of the themes or categories constructed from the findings except the researcher’s knowledge employed in uncovering themes or patterns of errors. The discovered themes are analysed and interpreted using critical realism stances using qualitative methodology to build categories and to form a questionnaire to reanalyse the scripts further using a quantitative methodology.

3.3.2.2 Critical Realism.

Critical realism paradigm acknowledges the existence of realities independent of human's perceptions; however, critical realism accepts human involvement in understanding the realities (Blaikie, 2007; Denzin & Lincoln, 2011; Evely, Fazey, Pinard, & Lambin, 2008; Guba & Lincoln, 1994; Maxwell, 2012; Moon & Blackman, 2014; Wong, Greenhalgh, Westhorp, & Pawson, 2012). Critical realism lessens the limitations posed by the two extreme research philosophies, positivism and relativism, by incorporating the realist ontology and the constructivist epistemology (Clark, MacIntyre, & Cruickshank, 2007; Denzin & Lincoln, 2011; Maxwell & Mittapalli, 2010); constructivist epistemology is associated with the interpretive paradigm. More information about interpretive philosophy is given in section 3.3.2.4.

Critical realism permits using the mixed methods methodology to investigate objects at different points. As reported by Roberts (2014) "critical realists argue that the world is "layered" into different domains of reality" (p. 2) ; and these domains of reality have different properties that can be studied using different approaches, qualitative and quantitative, and "if each is focused on different properties of social objects, then it is possible to reconcile them" (Scott, 2007, p. 10). Critical realism also accepts the phenomenon to be investigated as being concept dependent, requiring a human involvement which is the position shared by interpretivism (Sayer, 2000). On the other hand, critical realism acknowledges that causal explanation is also acceptable because a "material change in society has to be explained too" (Sayer, 2000, p. 18). By providing these perspectives, critical realism provides opportunities for both quantitative and qualitative researchers to achieve "a much needed systematic understanding of the relationships, structures and mechanisms constituting the material and social world" (Zachariadis, Scott, & Barrett, 2010, p. 20).

The analysis of student scripts has shown the presence of many different realities. Shallow and critical realist ontological assumptions have been taken to explore these scripts at different stages. First it was necessary to choose appropriate student workings (examination's scripts) from the complete set of such scripts. Second, from the selected scripts, it was necessary to identify whether these scripts contained valuable information that would assist the researcher in understanding the nature solutions of an MSQ further. Third, the shallow realist stance has been assumed to analyse the scripts because there were no specific research questions to apply to the analysis of student workings; therefore, the entire student workings were analysed to uncover any valuable information. Inductive qualitative content analysis has produced patterns of errors and solution strategies. Fourth, the critical realism position was adopted to evaluate the entities produced during the qualitative inductive analysis of the scripts, so as to transform the results into categories and subcategories; this activity requires the researcher's involvement. At this stage, errors and solutions strategies are new realities needing another investigation to uncover the relationships between them or to obtain further quantitative information from student workings to support qualitative results. The categories, subcategories and their descriptions were used to form variables which are part of the content of the questionnaire that was used as the quantitative phase instrument.

At level 1, in Figure 3.3, failing to distinguish the suitability of examinations scripts for containing appropriate information that fulfils the purpose of the research leads to the failure of the entire research project. At level 2, choosing the appropriate scripts ensures that the relevant information is uncovered. At level 3 the analysis of the selected scripts is used to form categories and subcategories.

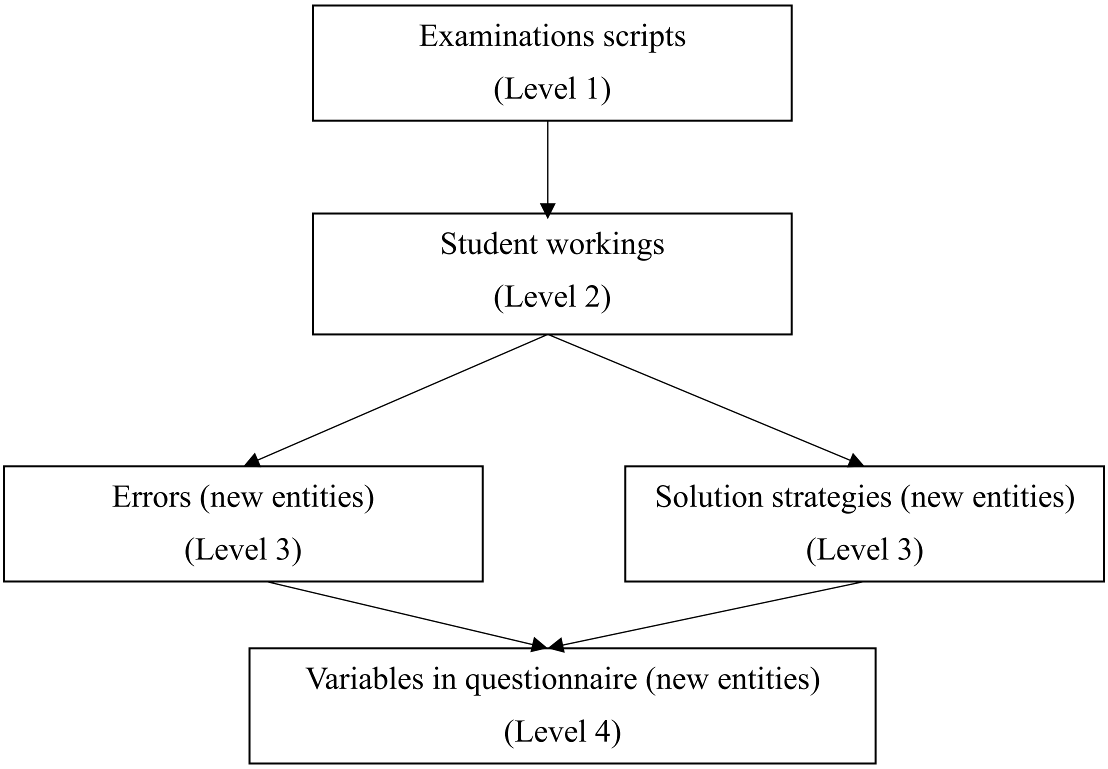


Figure 3.3 Layered ontological examinations scripts entities

At level 4, the variables in the questionnaire can only be formed from the outcome of studies that have been made using entities that reside at levels 1, 2 and 3. All these entities are ontological objects that are considered as layers of realities in which the application of multiple methodologies yields the knowledge that is regarded as relevant and useful to create knowledge that can be used in the domains of teaching, building and evaluating purposeful DSR artefacts. The research philosophies that enhance the understanding of selecting relevant methodologies are discussed in the following sections.

3.3.2.3 Relativism.

In relativism, which is also called “extreme subjectivism” (Evely et al., 2008, p. 52) and “idealist” (Blaikie, 2007, p. 16) ontology, subjective multiple varying realities are constructed by individuals, and as such, there is no independent realities that exist (Blaikie, 2007; Guba & Lincoln, 1994; Moon & Blackman, 2014). However, there are numbers of relativist ontologies with varying ontological positions. According to Blaikie, the idealist ontological positions vary from one extreme to the other; from the stance that fully rejects the existence of external world to the other position which considers "the construction of reality as just different ways of making sense of an external world" (p. 17).

In this research, since there are many stances along the continuum of the relativist ontology, the stance that recognises the existence of the external world can be applied in the analysis of examinations' scripts and in the development of an artefact. In the next section, the interpretive paradigm is discussed to determine if it offers better philosophical stances to be adopted in the current research activities.

3.3.2.4 Interpretive.

The word interpretivism is also known by many other terms; to mention a few, these terms are anti-positivist, constructivism, subjectivism, naturalistic, qualitative and humanistic (Hudson & Ozanne, 1988; Mack, 2010). These terms reflect the philosophical perspectives that the interpretivism paradigm represents. For example, the term subjectivism delineates the epistemological position of the interpretive paradigm, while the qualitative term indicates the methodology that interpretivist researchers generally use. The nature of "interpretive studies are, typically, inductive in nature and often associated with qualitative approaches to data gathering and analysis" (Gray, 2004, p. 37). The inductive qualitative analysis method was applied in the content analysis of student scripts at the qualitative phase. In qualitative research, interpretation of what is uncovered during research activities is essential.

According to Chen et al. (2011) and Ritchie et al. (2014), understanding or knowledge is gained through interpretation which involves researchers' reflections on what happened not only narrating the event, or observing without evaluation to make sense and meaning through "subjective interpretations of reality" (Chen et al., 2011, p. 129). This interpretivist's ontological stance differs from the ontological realist position that acknowledges the external existence of reality. However, although the concept of multiple realities is one of the relativist ontologies, “a number of interpretivists tend to stand closer to the realist side” (Chen et al., 2011, p. 133). Gray (2004) states that “while interpretivism and objectivism hold different epistemological positions, both are still based upon a *being* ontology (Chia, 2002)” (Gray, 2004, p. 23).

As discussed above, the ontological and epistemological stances of the interpretive paradigm show the relevancy of application of the interpretive philosophy to analyse student scripts. Morgan (2014) believes that

> any science, including a social science such as economics, ought to be about reality, the nature of reality ought to make a difference to how we study it and what we then claim about it as a form of knowledge. (p. 15)

In conclusion, scientific research makes inquiry about research objects that must have form or forms. Otherwise, there is nothing to investigate. The nature of these research objects and the ways they are perceived and studied determine the quality or kind of knowledge that is produced.

3.3.3 Epistemology.

In section 3.3.2, various types of ontological stances have been investigated and the most relevant stances for analysing the student scripts have been adopted. The adopted positions have identified types of realities and knowledge that can be revealed. Useful knowledge can be created in different ways using applicable epistemological stances. ‘Epistemology’ is the theory of knowledge containing assumptions that are employed in discovering, understanding, creating and formulating knowledge regarding the world or realities (Bernal, 2002; Blaikie, 2007; Moon & Blackman, 2014; Ritchie et al., 2014). Epistemology focuses on three major issues: (1) the most useful means of obtaining knowledge, (2) the relationships between inquirers and the inquired and (3) the issue of resolving the acceptability of the claimed knowledge as precise or real.

The first issue, which discusses the way of obtaining knowledge about the inquired object, as Ritchie et al. (2014) and Blaikie (2007) suggested, is to use one or multiple knowledge acquisition methods such as ‘inductive’, ‘deductive’, ‘retroductive’ and ‘abductive’. To acquire knowledge from the student scripts, inductive, deductive and abstraction methods have been used. Section 4.2 gives more detail about an inductive approach and a deductive approach. Section 4.4.4 gives more details about acquiring knowledge using the abstraction method. The deduction, reflection and abstraction methods have been applied to obtain the knowledge resulting from building and evaluating DSR artefacts. The knowledge acquisition methods are linked to the adopted ontological and epistemological stances. There are strong relationships between epistemology and methodology as Kivunja and Kuyini (2017) report that methodologies are applied to produce knowledge, which is the output of research efforts; these research efforts include the consideration of epistemology which “is concerned with the very bases of knowledge – its nature, and forms and how it can be acquired, and how it can be communicated to other human beings“(p .27). Thus, “epistemology can therefore be considered as an “intermediator” between ontology and methodology” (Aier & Fischer, 2011, p. 35).

The second issue involves the impact of relationships between researchers and the studied objects on the produced knowledge. Furthermore, Blaikie (2007); Creswell and Creswell (2013); Denzin and Lincoln (2011); Sobh and Perry (2006) agree that epistemology is concerned with the relationships between researchers and the researched realities. These relationships involve researchers’ views of the realities they investigate, which could be belief such as these realties existence independent of researchers’ thinking or constructed by them. This issue is associated with the questions of value free and value laden concepts. The topic of relationships between researchers and the studied objects lead to the discussion of the relationships between ontology and epistemology. As Crotty (1998) acknowledged “to talk of the construction of meaning is to talk of the construction of meaningful reality" (p. 10). Therefore, it is challenging to split the topics of epistemology and ontology. To acquire knowledge, the entities about which knowledge is sought must be there in any form, be it imaginary or real (Moon & Blackman, 2014). Guba and Lincoln (1994) link the type of relationships that researchers choose based on the adopted ontological stance(s).

The third issue, which is resolving the reliability of the knowledge from the inquired reality, is achieved in various ways. For example, in a social reality, the produced knowledge can be validated through consensus. The validation of the produced knowledge can also be accomplished by comparing it with previous similar research outputs; the other option is to perform several investigations of the same reality in which the outcomes of the investigations are expected to produce the same or similar outputs. Further detail is given in Ritchie et al. (2014, p. 8). In this research, the validity of the analyses of student scripts can be confirmed, for example, by comparing the uncovered patterns in one student workings with the other student workings in the same examination, and also across different student workings; the patterns in the findings can be verified by checking whether these patterns reveal the same or similar knowledge. All the scripts that have been analysed display similar characteristics because all the questions in different examinations test the same type of knowledge. Therefore, student workings from one semester's examination are similar to all other student workings; the only difference between the questions of the examinations is the parameters in the questions. Note, the knowledge that is required to produce solutions to the questions in all examinations is the same. Therefore, the results of the analyses are expected to contain similar patterns in all scripts. Sources of validation can also be external; this raw data is available for other researchers to perform analyses on them; and if the same criteria were applied to analyse the scripts for the same research purposes, the researcher believes, the same knowledge would be produced. Validation methods for the knowledge that will be produced from the analyses of the student scripts, will be shown in sections 5.1 and 5.4.2.

In summary, the production of knowledge, according to Crotty (1998), is through the way of looking at the world and making sense of it using research philosophies that include ontological and epistemological positions and using methodologies and methods that are influenced by the assumed stances. Table 3.1 shows the research philosophies that have been investigated including the ones that have been selected to use in this research.

Table 3.1 Current research philosophical positions

Paradigm	Ontology	Epistemology	Methodology
Positivism	Single reality and Independent	Objectivism	Quantitative
Interpretivism	Multiple reality (relativist ontology) independent reality (closer to the realist side)	Socially constructed (constructivism) subjective	Qualitative
Relativism	Multiple reality, independent reality	Relativist, subjective, constructionism	Qualitative and quantitative
Critical realism	Interpretivism and positivism	constructivism and objectivism	Qualitative and quantitative

The entities in Table 3.1 and the relationships between them are briefly described in the following paragraphs. The relativism and interpretivism paradigms have various ontological and epistemological stances. Along their continuum, they meet at some points where some of their ontological views accept multiple or independent realities.

Epistemologically they furthermore agree that subjective view can be applied when investigating real or mind-formed realities. In relativism paradigm, epistemological relativism and/or constructivism stance(s) can also be exerted. In the interpretivism paradigm, constructivism epistemological position can also be applied to multiple and single external realities. Critical realism paradigm includes the mentioned ontological (interpretivism and positivism) and epistemological (constructivism and objectivism) assumptions that researcher has chosen to select the relevant methodologies.

Student workings in the scripts are independent entities and the knowledge is created constructively at the qualitative level. Further knowledge is created objectively at a quantitative level. Hence, the researcher had concluded that the interpretivism ontological view of independent realities and constructivist epistemology, positivist ontological and objectivist epistemological views present better options to conduct the analyses of student scripts to produce acceptable and purposeful knowledge. In the next section, the discourse is about the ontological and epistemological stances of DSR artefacts.

3.3.4 Ontology and epistemology of artificial science.

In IS research as Ken et al. (2007) and Carlsson, S.A. 2006 (as cited in Alturki, Bandra & Gable, 2012) state the paradigms that are familiar in IS research are social and behavioural science research. The IS research paradigms, according to Galliers 1991 (as cited in Venable, 2006), Burrell & Morgan 1979; Chen & Hirschheim 2004; Fitzgerald & Howcroft 1998; Hirschheim & Klein 1989; Iivari et al. 1998; Lee 1991; Monod 2003; Weber 2004 (as cited in Niehaves, 2007) and Orlikowski and Baroudi 1991 (as cited in Korpela et al., 2004) are positivist and interpretivist; furthermore, Venable (2006) and Orlikowski and Baroudi (1991) add the usage of critical realist paradigm as being the third paradigm for IS research. The paradigms for DSR in IS are the same as the IS research paradigms, because the purposes of the two research objectives, in this case, are the same. Further details about the IS research paradigms were given in the discussions that have been made in sections 3.3.1 and 3.3.2.

Iivari (2007) divides the ontology in DS into three “worlds”. World 1 involves evaluation of IT artefacts in which the research phenomena involves the nature of world and IT artefacts. World 2 deals with the evaluation of IT artefacts in relation to the “perceptions, consciousness and mental” (Iivari, 2007, p. 42) world. World 3 consists of research phenomena that involves IT artefacts, institution, theories, IT application, evaluation of institutions’ information systems. World 3 could also provide the possibility of IT artefacts contributing to the creation of new theories. In this research, the DS ontology fits in the world 3; the literature review chapter has evaluated the nature of ICT assessment tools to understand how these tools assess subjective questions; the results of the evaluation has revealed that the ICT mathematical assessment tools face challenges to assess MSQ solutions. The creation and evaluation of artefacts in this research also fits into the DS’ research phenomenon of world 3. These artefacts, once they are built, are considered to be real world beings.

Vaishnavi, Kuechler, and Petter (2017) elaborate on the metaphysical position of DSR artefacts as being unique entities. Vaishnavi, Kuechler, and Petter give two reasons for assuming the uniqueness of DSR artefacts’ existence. One of the reasons are the changing of these assumptions as the artefacts development and evaluation goes through different stages during the life of the research. However, Vaishnavi, Kuechler, and Petter continue their arguments stating that the ontological nature of DSR artefacts as being one real world.

Bunge (1979) classifies technologists as the critical realists with tendency towards pragmatic attitudes. Reality in view of technologists is the “sum total of resources (natural and human)” (Bunge, 1979, p. 194). Gregg, Kulkarni, and Vinzé (2001) take the position of reality in DSR as being an object in a known context that is created technologically and socially. All these authors view DSR artefacts as real objects that display the characteristics of one world. However, they are different realities at different stages, but not multiple realities at the same stage. In this research, the knowledge derived from the analyses of scripts, philosophical assumptions taken, the researcher efforts, justificatory knowledge, kernel theory are these resources that technologists refer to. These are realities in different forms utilised to extract the knowledge that technologists require to achieve their goals.

According to Vaishnavi and Kuechler (2015), the sources of knowledge in DSR outputs are (1) the knowledge that is used to develop and evaluate the DSR artefacts and (2) the functionality of the developed DSR artefacts. The artefacts are characterised in the statement "what it means is what it does" (p. 22) regarding their functionality. For Bunge (1979), technologists use the knowledge that comes from the existing realities to achieve their goals. In DSR, this knowledge can be created by adopting objective and interpretive epistemological assumptions (Gregg et al., 2001). On the other hand, according to Vaishnavi et al. (2017), knowledge is acquired during artefact development using cognitive processes that involve abduction, deduction and circumscription methods. Circumscription "is a rule of conjecture that can be used by a person or program for 'jumping to certain conclusions'" (McCarthy, 1980, p. 2). Researchers learn or discover new knowledge when a circumscription process is applied during the artefacts' construction and evaluation cycle.

Researchers' relationships with the artefacts under development is both objective and interpretive; they discover new knowledge and/or constraints during the artefacts 'build and evaluate' cycles. The discovered knowledge and/or constraints occur(s) objectively without involvement of researchers. However, when researchers utilise the discovered knowledge and/or resolve the constraints, researchers' roles change from spectators to the active participants which is interpretive. These assumptions are the same as that taken to analyse the student scripts. Thus, the type of research phenomena determines the ontological and epistemological stances that are taken in IS research and the stances taken influence the research methodology to be used in the IS research (Iivari, 2007). This status leads to the discussion of the identification of research methodology that is explained in section 3.3.5.

3.3.5 Methodology.

Research methodology is an important research tool that researchers need to carefully construct to achieve their purposes. Some researchers refer to the words qualitative and quantitative as method, methodology and paradigm. However, in this research, these terms have different meanings and are used for different purposes. The definition of the word paradigm has been given in section 3.3.1 while the definition of the word methods will be given in 3.3.7. The term methodology is used when referring to both qualitative and quantitative words. It has many different meanings; methodology "is a structured set of guidelines or activities to assist in generating valid and reliable research results" (Mingers, 2001, p. 242). Bryman (2008), Creswell and Creswell (2013), Holden and Lynch (2004) and Leedy and Ormrod (2012) identify methodology as a general approach or "all the means available" (Holden & Lynch, 2004, p. 6) that researchers employ to perform research when studying phenomena.

It is important to consider the differences between a research process and a research methodology. The research process includes the entire research activities including the communication of the research. Nunamaker et al.'s (1990) 'research process' definition has been given in section 3.2. Methodology in the context of research "consists of various methods or techniques" (Mingers, 2001, p. 242). Qualitative and quantitative methodologies have research designs that include research methods. In the following sections qualitative, quantitative and mixed methods methodologies will be explained.

3.3.5.1 Qualitative Methodology.

There is no consensus on the definition of qualitative methodology (Guest, Namey, & Mitchell, 2013). Qualitative research methodology (Fetters, Curry, & Creswell, 2013) is employed to analyse, explore or unpack (Astroth, 2018; Bansal, Smith, & Vaara, 2018; Fetters et al., 2013; Venkatesh, Brown, & Bala, 2013) a phenomenon (Astroth, 2018; Fetters et al., 2013; Venkatesh et al., 2013) or to answer questions or explain the studied objects (Astroth, 2018) in order to produce deep understanding and/or to create a theory (Fetters et al., 2013; Venkatesh et al., 2013). Qualitative methodology presents researchers with many advantages. The nature of qualitative methodology being inductive and flexible gives choice to researchers, for example, to change "the sampling procedures during the data collection process based on incoming data" (Guest et al., 2013, p. 4). Furthermore, in qualitative research, researchers can probe for more information using open-ended type questions aiding researchers to reveal reliable research outcomes (Guest et al., 2013). The general qualitative research designs and the design for this research to analyse student scripts will be discussed in sections 4.1 to 4.4 inclusive.

3.3.5.2 Quantitative Methodology.

Quantitative research methodology is concerned with finding the relationships between concepts or variables to test theories objectively (Creswell, 2014; Newman, Ridenour, Newman, & De Marco, 2002) and/or to find answers to "research questions about causality, generalizability, or magnitude of effects" (Fetters et al., 2013, p. 2). Various types of quantitative research designs and the design for analysing scripts quantitatively will be discussed in section 4.5 and section 4.6. In these designs, quantitative research might require qualitative research data. O'Cathain, Murphy, and Nicholl (as cited in Fetters, et al., 2013) identify the usage of qualitative research data to form quantitative research instruments and to form hypotheses to be tested. Fetters et al. (2013) report that some qualitative research extends the research to include quantitative research to better explain the studied objects or uncover new knowledge concerning the studied objects. In this research, the qualitative phase of the analysis of student scripts will be used to form the questionnaire instrument to be used in quantitative phase. The result of the quantitative phase will be used to support and explain the results of the analysis of student scripts performed at the qualitative phase. The existence of the relationships between the two research methodologies indicate the existence of another methodology which is known as mixed methods.

3.3.5.3 Mixed methods methodology.

Researchers use mixed methods, which is based on both qualitative and quantitative methodologies, (adore Newman, Ridenour, Newman, & DeMarco Jr, 2003; Ågerfalk, 2013; Johnson & Onwuegbuzie, 2004; Newman et al., 2002; Venkatesh et al., 2013) to capture "the strengths of both quantitative and qualitative approaches" (Fetters et al., 2013, p. 2) and for the ""better understanding" of the inherent complexities and contingencies of human phenomena" (Greene & Hall, 2010, p. 2). The qualitative and quantitative methodologies originate in different paradigms. As discussed in section 3.3.2.2, critical realist stances were adopted to analyse the student scripts. Critical realism allows mixing methodologies leading to the formation of a pragmatist paradigm with the view that "the methodology is adequate if it leads to the solution of the research question" (Mayring, 2014, p. 6).

Regarding mixed methods research designs, usually there are two types of designs which are sequential or concurrent. These designs are further subdivided to form more design types based on the priority that is given to the selected research design phase. For example, if the sequential design were applied beginning with the qualitative phase followed by the quantitative phase, the researcher could give the priority to either of the phases or could give them equal weight. More details about mixed methods research design are given in Hanson, Creswell, Clark, Petska, and Creswell (2005, p. 227).

To analyse the student scripts, the researcher used a sequential design starting with the qualitative phase that was followed by the quantitative phase and this design is called ‘exploratory mixed’ methods. The purpose for the philosophical assumptions behind using the exploratory design is explained by Creswell and Clark (2017:

> Since the exploratory design begins qualitatively, the research problem and purpose often call for the qualitative strand to have greater priority within the design.
>
> Therefore, researchers generally work from constructivist principles during the first phase of the study to value multiple perspectives and deeper understanding. When the researcher moves to the quantitative phase, the underlying assumptions may shift to those of postpositivism to guide the need for identifying and measuring variables and statistical trends. Thus, multiple worldviews are used in this design, and the worldviews shift from one phase to the other phase. (p. 87)

The purposes for using qualitative methodology at the initial step has been explained in section 3.3.2.2. The priority was given to the qualitative study, so that using “multiple perspectives and deeper understanding” (Creswell & Clark, 2017, p. 87) will help to produce better outcome from the analysis of student scripts. The outcome of this research phase was categorised, and concepts were formed. Then the categories and concepts were used to establish the quantitative research phase. The quantitative process was used to analyse the student scripts further; the activities performed during the quantitative research phase alerted the researcher to do the reviewing and refining of the initial qualitative research phase’s output; this was achieved by doing further analysis of student scripts qualitatively. The advantages of using multiple methodologies to acquire knowledge is epistemological issue producing valid knowledge that has been discussed in section 3.3.3.

3.3.6 Axiology.

Axiology is the research philosophy that deals with values researchers bring to their research as well as the roles of these values in research (Biedenbach & Jacobsson, 2016; Saunders, Lewis, & Thornhill, 2016; Tashakkori & Teddlie, 2010); Saunders et al. (2016) add ethics besides values entity to form axiology, or simply saying axiology is what researchers bring with them to the research as a form of belief (Creswell & Creswell, 2013). In research philosophy, adding the aspects of axiology to ontology, epistemology and methodology strengthens the validity of the research that consequently enhances the acceptability of the knowledge contributed by the research (Aldawod & Day, 2017).

Researchers' responsibility is to consider the results of their efforts and effects on other beings which are physical and/or abstracts (Iivari, 2007). This is to indicate that researchers have situated themselves in regard to their view at all stage of the research process. Pham (2016) denotes, in value-laden research, that researchers own values might interfere with the entire research activities; therefore, this type of research should be done with great attention. The type of values held by researchers varies from discipline to discipline. Iivari's (2007) view is that design science research cannot be value free; consequently, design science researchers must declare their positions explicitly.

The value-laden and value free- stances are present in the scripts' analyses, and the declaration of the researcher's values reveals the type of knowledge that is contributed regarding the means that was used as basic knowledge to build and evaluate the artefacts.

The building and evaluation of the artefacts process was performed objectively, since as explained, DSR cannot be value laden. However, although this stance is declared, it is an inevitable to involve the researcher's role or values in the creation of artefacts as circumscription process includes both positions which are value laden and value free positions. Nevertheless, the two values are not at the extreme ends of the values assumed allowing the researcher to be creative and innovative during the creation of DSR artefacts. The definition of axiology has included the research ethic. The researcher had sought permission from Victoria University to get access to the scripts.

These scripts were considered as the properties that belong to the students who participate in the examinations. However, since the university is the custodian of the scripts, they have the authority to permit others subjects to get access to them. Following this understanding, the researcher, had applied to access these scripts and received approval to use the scripts for the research purpose.

The researcher copied scripts after removing students' identifications and used the copied materials as the data source. After selecting the relevant research philosophies and methodology, section 3.3.7 discusses research methods.

3.3.7 Methods.

Researchers are required to describe and justify the research methods to be used and act upon the outputs and consequences resulting from the methods' application. Creswell (2003) and Crotty (1998) specify that research methods are specific descriptions of procedures that contain the activities involving the collection of data to be analysed.

Methods "are generally well-defined sequences of operations that if carried out proficiently yield predictable results (Checkland 1981, Livari et al. 1998)" (Mingers, 2001, pp. 241-242). They can be used, for example, "to describe phenomena, to observe their interrelationships, and to make predictions about those interrelationships" (Riff, Lacy, & Fico, 2014, p. 32).

Crotty specifies the need to illustrate type of methods, the ways these methods are used, and the context and the setting that these methods are employed in. Furthermore, Crotty underlines the issue of depicting the type of actual activities, such as (1) uncovering themes and recognising or interpreting their meanings, (2) how these themes arise and the means they are recognised by, and (3) the actual actions taken regarding the emergent themes. Thus, from the information thus revealed, the necessity to show the details of activities in research methods is specific to the particular research conducted; however, when developing these methods, it is important to remember their relations with other research philosophies such as methodologies; because the type of methodology adopted influences the determination of the overall nature of the details of activities in the research methods. The suggestions that have been made by Crotty will be implemented in chapter four.

3.4 Summary

The focal points of this chapter have been (1) to examine the associations between natural, social and design sciences, (2) to review selected DSR frameworks and create a DSR framework for this research and (3) to investigate the research philosophies to reveal the relevant research philosophies for this research. The results of the investigation of the relationships between the three branches of sciences have shown the need of design science research for the knowledge that is produced by the other research outputs. The investigation further establishes that DSR needs both design and other sciences' research methodologies.

The DSR frameworks' review and examination has resulted in creating a DSR framework named DSRA-MM. The content of this framework has been shown in Figure 3.2. The sources and steps that have been used to construct the framework have been detailed in section 3.2. The framework consists of six phases and will be used to guide the researcher throughout this research.

The tasks in the first phase of the DSRA-MM framework have been explained and implemented in chapter two. The second phase of the framework consists of two main consecutive parts. The first part was about investigating the research philosophies. This investigation has included the discussions of the natures of the research paradigms, ontologies, epistemologies, methodologies, axiology and the research methods. The outputs of the review have displayed the importance of researchers understanding these philosophies and adopting their research relevant philosophical stances. Furthermore, the outputs disclosed the influences of researchers' stances on the entire research efforts.

Based on the outcomes of the philosophies investigation the critical realism philosophy has been adopted to select the relevant methodologies to analyse the student scripts and build DSR artefacts. The investigation of the methodologies has resulted in establishing the relevant methodology to analyse the scripts as being the sequential exploratory mixed methods methodology. Additionally, the review outputs recognised the usage of design science method to construct DSR artefacts. The usage and/or implementation of the selected methodologies will be shown in the subsequent chapters. The tasks in the second part of the second phase in the DSRA-MM framework are to analyse the student scripts; this topic will be covered in chapter four. The development and building process of DSR artefacts will be covered in chapter six.

Chapter 4 Analyses of student scripts

In chapter three, various research philosophies have been discussed, generally, regarding the student scripts and DSR artefacts. The philosophical stances regarding these two entities have been determined; the selected stances have been used to select the methodologies to investigate the student scripts and to develop the artefacts.

In this chapter, the focus is on analysing the scripts using a sequential exploratory mixed methods methodology. The qualitative and the quantitative methodologies' research designs will be investigated in order to select the relevant research methods.

This chapter is divided into several sections. Various qualitative designs to collect, analyse and report data belonging to the researched objects are mentioned in section 4.1; the section also states the design that is selected to analyse the student scripts qualitatively. The selected design and its implementation are identified in section 4.2. The section describes the types of content to which content analysis design can be applied. This includes written documents as the source of the researched objects from which meaningful data can be drawn. This section, further, examines the application of an inductive approach to the content analysis design and the justification that has been made to use this approach to analyse the scripts during the qualitative phase. The source where the scripts are originated in and the educational level of students whom scripts belong to are briefly stated in section 4.3. Section 4.4 and its subsections investigate various techniques that are suggested by different authors to analyse the content of the student scripts. This section also examines in detail the technique that is adopted to analyse the scripts. This technique consists of five steps: (1) sampling, (2) collecting and managing raw data, (3) preparing raw data, (4) organising data and (5) reporting the result of analysis. The first step, sampling, is explained in section 4.4.1. This section explores the level of the analysis that might include the depth of analysis and the scope of the collected data sources for the purpose of determining what to analyse. The sampling process itself consists of seven steps and is explained in the subsections of the sampling section.

Section 4.4.1.1 involves the first and the second steps of sampling. This section determines the target population and explains why identifying sampling frame is difficult. The section details (1) the reason for determining the target population, (2) the definition of a target population and (3) the reason for conducting research on a sample population instead of the target population. The justification for considering the student scripts as a sample of a target population is also explained in the section. The third step in a sampling process is determining the sampling method which is discussed in section 4.4.1.2. The definitions of a sample and sampling are given in that section followed by a discussion about the general types of sampling methods which are probability and non-probability. Quantitative research, generally, uses a non-probability sampling method. The section investigates the most appropriate method that can be used to select a sample of a target population. The rationales for a sample to qualify as the sample of the target population are discussed and the purposive sampling method is selected as the appropriate sampling method for both qualitative and quantitative phases. Finally, the section concludes by presenting the justification that is used to select the purposive sampling method to choose the student scripts as the sample of the target population.

Using the purposive sampling method, the scripts are scanned to select the appropriate scripts that fulfil the criteria that are used to identify a sampling unit. The different definitions of the sampling unit and the roles of these definitions in complementing one another are the focuses of section 4.4.1.3. Furthermore, this section explains the criteria that are used to assist in identifying a student script to be considered as the sampling unit. In the sampling process, after the identification of the sampling unit, the other important task is to determine the sample size; this issue is explained in section 4.4.1.4.

This section presents ways that assist in establishing the sample size that can adequately represent the target population when conducting qualitative and quantitative research. Three factors in particular are discussed to determine the size of the sample in quantitative research. For qualitative research, the section discusses how to reach the saturation level to obtain the sample size; furthermore, the section discusses how the saturation level is reached and presents the argument to justify that the number of collected scripts are sufficient to produce enough data to achieve the research purpose. Finally, the last step in the sampling process is to detail the implementation and reporting of the sampling plan and these issues are discussed in section 4.4.1.5.

The steps that are used to collect and manage the student scripts are explained in section 4.4.2. In this section, the researcher's familiarisation with the scripts, sampling unit selection criteria and the results of applying the criteria will be discussed. Additionally, steps that are adopted to manage the collected scripts will be detailed. Once the scripts are collected and managed, the next task, which is preparing the scripts for analysis, will be explained in section 4.4.3.

In section 4.4.4, the data in the prepared scripts is organised. The act of organising data involves open coding. Multiple levels of coding methods are used to deeply explore the scripts' content to help comprehending the student workings in the scripts. The depth of analysis creates deep reflections about the data that to be analysed and the data that is extracted from the scripts. The relevant and useful thoughts can be captured in a memo. The section reports a memoing technique and how it is used to acquire those thoughts and reflections. Furthermore, this section explains steps and tools that are used in analysing the scripts. In the end, the section details the qualitative output of the analysis of student scripts. The next sections of the chapter relate to the quantitative analysis of student scripts.

Multiple designs of quantitative methodology are examined in section 4.5. These examinations include experimental and non-experimental types of designs. The detailed purpose of a survey design and its types are described in section 4.5.1. The two types of survey design that are descriptive and analytical are explained. To implement the selected type of survey design, section 4.5.2 identifies the deductive approach. The enquiry into the types of instruments to collect the data from scripts in a survey design is discussed in the section. The process that is followed to form a new survey questionnaire is discussed in section 4.5.3.

The main types of data to be collected from the student workings, the purposes of collecting this data and the relationships between variables to be measured, the questions in the questionnaire and the research questions are discussed in section 4.5.4. At this stage the collected data is ready for analysis. To enter the data into tables the data should be in the format that enables the selected software to analyse it. The creation of the data matrix and the details of its rows and columns are reported in section 4.5.5.1. The exploration of the data in the tables is covered in section 4.5.5.2. This section describes the actions to be taken to correct the explored data, before the final analysis is performed. Section 4.6 concludes this chapter.

4.1 Qualitative Student Scripts Analysis Approaches

The purpose of the qualitative phase, in this research, is to analyse the content of student scripts for the purpose of extracting data such as solution strategies, student errors and awarded marks. The qualitative methodology has different research designs allowing the selection of designs that are relevant for research. Guest et al. (2013) identify participant observation, in-depth interviews, focus groups, and document analysis as being the most prevalent types of designs in qualitative research. Guest et al. (2013) and Creswell (2014) state the existence of many approaches to collect, analyse and report qualitative data. Guest et al. believe that the most popular approaches in the qualitative research, regardless of fields, are phenomenology, ethnography, inductive thematic analysis, grounded theory, case study, discourse/conversation analysis, narrative analysis and mixed methods. Rose (2015) presents the same approaches and adds action research and language analysis designs which include qualitative content analysis, conversation analysis and discourse analysis. After reviewing of all these designs the content analysis type has been determined as the most relevant design to analyse the students.

4.2 Content Analysis and Inductive Approach

Content analysis design, which is also called content analysis method, involves the process by which research relevant data in the content of an object are identified, collected and classified based on determined features after which the conclusion is drawn in regard to the analysed content (Rose, 2015). The word content refers to "what is contained …. in a message" (Prasad, 2008, p. 2). Usually, the content analysis method is used, in quantitative and qualitative research, to analyse data that is obtained from human interaction processes, verbal or visual, and written documents (Elo & Kyngäs, 2008; Forman & Damschroder, 2008; Mayring, 2000). Creswell (2012) emphasises the importance of documents as being one of "a valuable source of information in qualitative research" (p. 23). Creswell details further the role of documents providing "information in helping researchers understand central phenomena in qualitative studies" (Creswell, 2012, p. 223). Student working in the scripts is the central phenomenon in revealing information. The content analysis method can also be applied to all contents that contain meaningful data because it is "a method used to answer research questions about content" (Riff et al., 2014, p. 11). Many disciplines that seek to investigate contents in phenomena have been applying this method. For example, in the IS discipline, Al-Debei and Avison (2010) used content analysis in their study of business model concept in the field of IS using IS literatures while Arnott and Pervan (2012) explored the decision support systems employing content analysis, to uncover the application of DSR research paradigm. The application of the content analysis design by different disciplines indicates its relevancy to analyse the content of the student scripts for the purpose of understanding the research problem and presenting information that assists in building and evaluating DSR artefacts.

Inductive and deductive approaches are used when using content analysis method. Both approaches can be used together or separately in one research. The deductive approach is covered in section 4.5. In the inductive approach, the starting point is exploring the data in a selected phenomenon in order to answer research problem(s) and/or to construct theory or category (Rose, 2015). In this research the purpose is to construct categories from the findings of the qualitative analysis of the scripts. This approach allows researchers using qualitative methodology to investigate research objects and draw knowledge from the collected data (Crotty, 1998; Hathaway, 1995; Moon & Blackman, 2014; Newman et al., 2002); Creswell (2013) claims no previous knowledge is required when using the inductive approach in an investigation; although Blaikie (as cited in Ritchie, Lewis, Nicholls, & Ormston, 2013) argues that the nature of qualitative research is not purely inductive because Ritchie, Lewis, Nicholls, and Ormston (2013) claim

> For example, when so called inductive researchers generate and interpret their data, they cannot approach this with a blank mind. Even if they are not testing a hypothesis, the kind of data they have generated, the questions they have asked and the analytical categories they have employed will have been influenced by assumptions deductively derived from previous work in their field (p. 6)

In the inductive approach, the view is that there is no ‘previously structured knowledge’, but scattered knowledge that is sufficient to pose questions about phenomena and background knowledge. This background knowledge enables the initiation of an induction process that involves identifying, collecting relevant data, and assisting in interpreting and drawing conclusion. This indicates that the inductive approach is applicable “if there is not enough former knowledge about the phenomenon or if this knowledge is fragmented (Lauri & Kynga¨s 2005)” (Elo & Kyngäs, 2008, p. 108). Hence, the inductive approach is appropriate for investigating unknown phenomenon or phenomena for which previous knowledge is scarce. There is no, to the knowledge of the researcher, previous research that can be used to guide the study of the student scripts. Thus, the discussions that have been made so far, regarding the qualitative content analysis design and inductive approach, make clear why the researcher has resolved to selecting this design to analyse student scripts. Before the application of the content analysis design, a brief description of the selected student scripts is given in the following section.

4.3 Student scripts – Investigated Phenomenon Description

The student scripts belong to a foundational first year mathematics unit that was taught to engineering students at Victoria University, Melbourne, Australia; the selected questions came from a linear algebra topic. The instruction that was given to the students, was to use Gaussian elimination to solve system of linear equations. The details, such as number of examinations, number of scripts in each examination for the selected questions are presented in section 4.4.2. The type of questions to be analysed are multi-steps questions (MSQ). These MSQs are considered similar to solving complex problem when ICT assessment tools are used to evaluate student’s complete workings to award marks. The qualitative content analysis steps that are used in analysing the student scripts are discussed in the next section.

4.4 Inductive Qualitative Content Analysis of Student Scripts

For conducting the inductive qualitative content analysis, many authors provide similar steps that generate required outputs. Mayring (2000), Mayring (2014), Elo and

Kyngäs (2008) and Rose (2015) suggest parallel procedures when using the inductive qualitative content analysis design to analyse phenomena. Rose, Elo and Kyngäs's steps are comparable, and the steps in Figure 4.1 are adopted from the combinations of their steps.

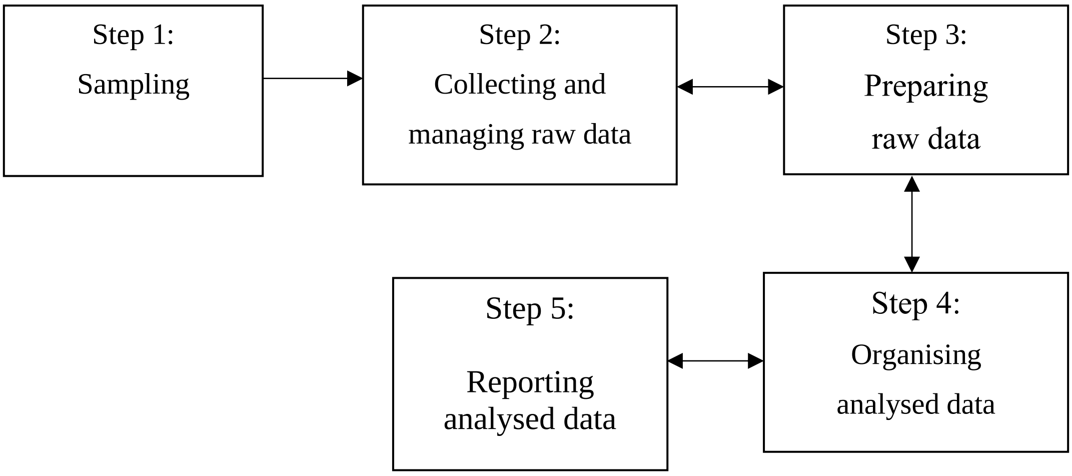


Figure 4.1 Inductive qualitative content analysis of student working steps

Figure 4.1 shows the process followed in analysing the student scripts in the qualitative phase. This phase is made up of five different steps. The first three steps relate to collecting of the raw data while the last two relate to the analysis of raw data. The outcome of the sampling process in Figure 4.1 is common for both quantitative and quantitative phases because the same research objects, which are student workings, will be analysed qualitatively and quantitatively but for different purposes.

4.4.1 Sampling.

When applying a content analysis design, the first concern in the preparation or sampling phase is to determine what to analyse; and this depends on the research problems and purposes. Once the purpose of the analysis is determined, it is important to decide the level of analysis and sampling issues. Elo and Kyngäs (2008) state that "deciding on what to analyse in what detail and sampling considerations are important factors before selecting the unit of analysis (Cavanagh 1997)" (p. 110). The level of analysis refers to the latent and manifest types in the content destined for analysis and "there has been some debate as to whether hidden meanings found in documents can be analysed, because their analysis usually involves interpretation" (Elo & Kyngäs, 2008, p. 110). In studying the student scripts, both types of content will be considered; the interpretation of the findings in the analysis is required to understand reasons why student workings produce specific patterns or information and the interpretation process enables building categories from the findings. The level of analysis also refers to the scope of the collected data sources which might include for example, individuals, groups and organisations (Rose, 2015).

The sampling matters that are associated with the selection of the unit of analysis are discussed next. Sampling involves series of steps, and the details of five critical steps are given in Taylor (2017, p. 165). In analysing the student workings, the sampling steps which are shown in Figure 4.2 will be implemented. These steps have been modified from Rose's (2015) sampling steps. The figure shows seven steps that are covered in the subsequent sections.

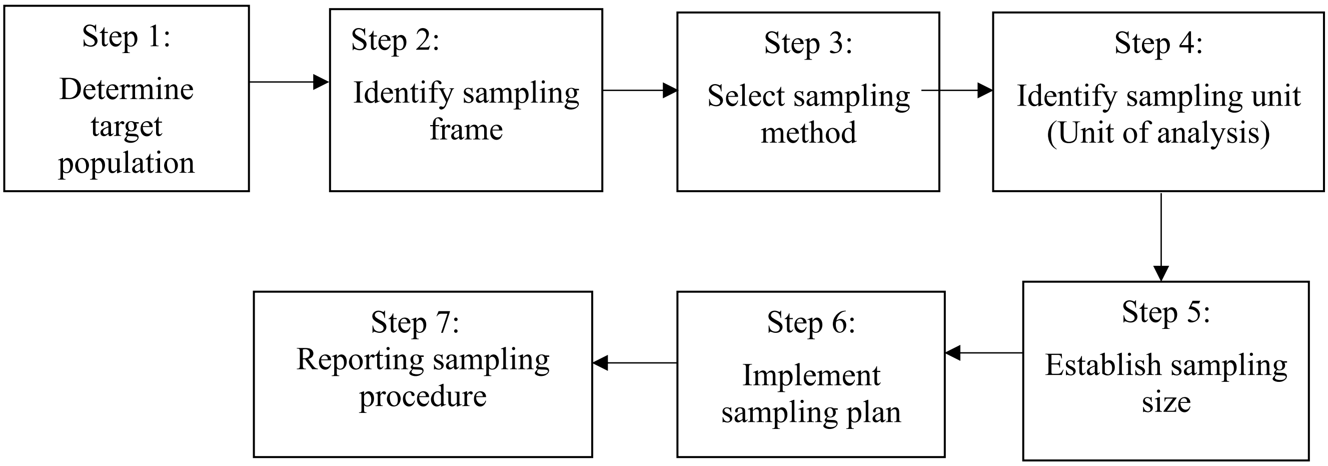


Figure 4.2 Student works sampling process - adapted from Rose (2015, p. 189)

4.4.1.1 Determining target population and identifying sample frame.

The following discussion concerns the first and second steps which are determining the target population and identifying the sampling frame. The first step, which is the identification and determination of the phenomenon- target population-, helps to establish the sampling methods and the sampling size. A population that is also termed a 'target population' or group includes the entire identified entities that could be abstract or physical, such as group of people and events, that are selected for observation (Lehman, O'Rourke, Hatcher, & Stepanski, 2013; Porta, 2014); a target population could be subjects "about which inferences are desired"(Porta, 2014, p. 279); furthermore, a target population is not limited to a particular group, but "often includes not only all persons living at present but also those that may be alive at some time in the future" (Kirkwood & Sterne, 2010, p. 10). Conducting a study on unlimited or large size of target population is impossible or very challenging; therefore, the alternative option is studying a sample of the population.

In this research, all student scripts that will be selected for analyses are considered the sample of the population; because the real target population is not only the students whose examinations' solutions are assessed in this research. It also includes other students who seek the same knowledge and take similar examinations to reveal the knowledge these student workings reveal. Therefore, it is impossible to identify a sampling frame "which is a listing of all the elements of … target population from which the sample can be drawn" (Rose, 2015, p. 189). The difficulty of identifying the sampling frame has led to the exclusion of 'identifying the sample frame' step. The next step after deciding the sample to be studied is to select the relevant sampling method.

4.4.1.2 Determining Sampling method.

A 'sample' is a subcategory of a whole population that is carefully selected with the intention of reflecting the characteristics of the population (Lehman et al., 2013; Riff et al., 2014; Rose, 2015). The selection of an appropriate sample from a specific population requires a process known as 'sampling' (Porta, 2014); sampling involves "deciding where you will collect your data and how much you will need to collect" (Rose, 2015, p. 187) for the purpose of conducting a study (Burns & Grove, 2010); and generally there are two types of sampling methods, which are named non-probability and probability.

In the non-probability sampling method, defined appropriate criteria guide researchers to select samples for observation. The main reason for using non-probability sampling, in quantitative research, is due to the absence of "an adequate sampling frame" (Riff et al., 2014, p. 74). The probability sampling has various types of sampling methods. However, in all these methods, the sampling units are randomly selected from the entire population giving the selection opportunity for all units in the population (Rose, 2015); many authors claim this process permits the sample to be a representative of a population. However, there is disagreement regarding the randomisation to be the only accepted method for selecting a sample that can be considered a representative of a population. The view that is taken in this research is "a study group can potentially be representative of a given target population, and its findings generalizable to it, even if the target population is not amenable to sampling" (Alexander, 2015, p. 3). Alexander continues this view by remarking that "a sample can be representative of a population if it is "typical in respect of certain characteristics, however chosen,"(p. 3); furthermore, the random selection of studied entities does not guarantee that the data source to be information rich (Palinkas et al., 2015). This argument shows the possibility of using the non-probability sampling methods to select the samples that can be representative of the population.

One of the non-probability sampling methods is purposive sampling, which is a sample selection method that is based on reasoning that is regarded purposeful in achieving the need of research (Riff et al., 2014; Rose, 2015); and it is the easy technique to select a sample that can represent the population, if researchers have prior in-depth knowledge about the phenomenon to be studied (Madsen, 2011). Madsen identifies, that by using the purposive method, "if we are lucky (and clever!) we can thereby obtain a sample, which is very similar to the population" (P. 108). One of the other intentions of applying this method is to "learn a great deal about issues of central importance to the purpose of the research" (Patton, 2002, p. 46). Hence, although this sampling method is generally used in qualitative research, it is also validly applicable to quantitative research following the arguments presented by Alexander and Madsen. Thus, the output of the qualitative content analysis using purposive sampling method can be used to analyse student workings quantitatively. The justification for using purposive sampling, when choosing the examinations' scripts, is based on the purpose of the research which is to uncover students' actions in solving questions given in the examinations to display their knowledge. The researcher has prior in-depth knowledge of the phenomenon which is the knowledge that students are requested to demonstrate through taking examinations.

4.4.1.3 Identifying sampling unit (Unit of analysis).

A unit of analysis, also known as the sampling unit, refers to the studied object; the unit of analysis has been given various related definitions. According to Mayring (2014), the unit of analysis is the segmentation rules that are used to divide the whole text into segments. Riff et al. (2014) regard it as a segment that is drawn from a whole content or object and it has the sufficient meaning that is relevant for observation to answer "one or more variables of theoretic interest" (Riff et al., 2014, p. 60) and the result of the observation is quantifiable.

The unit of analysis "may be the same as an individual in the population. It may also be a group of individuals from the population" (Madsen, 2011, p. 101). It is also defined as the degree of the combination of data when analysis is performed and when the presented findings are referred to as the unit of analysis (Rose, 2015). All these explanations complement one another regarding the position of unit of analysis in the content analysis process; furthermore, it is important to select the appropriate unit of analysis before analysing contents. According to Elo et al. (2014), the utmost acceptable unit of analysis is required to have the meaning that can form one complete element with the manageable size.

Each selected script containing a student working is measured against the criteria that: (1) it is the complete element which contains observable data that can be used in answering the research questions and (2) its size is manageable and can be judged as an entity representing a student, who is an individual in the sample of the target population. Hence, it is important to specify the factors that qualify the selection of the scripts as the units of analysis.

Various factors influence the selection of a unit analysis. The unit of analysis in the script analysis is determined by the type of the knowledge that scripts contain; this knowledge is a student working that contains solutions to multi-steps questions. These solutions must show at least some degree of understanding in solving these questions. The selection criteria and the researcher's domain knowledge ensure the appropriate collection of the data source to be studied.

4.4.1.4 Determining sample size.

The selection of the relevant research applicable population, the availability of the sampling frame, the sample selection criteria, the sampling methods, and the sample size are among important factors that determine the quality of data to be analysed. The relevance of the examinations' scripts as the appropriate population, the criteria for selecting them as well as the sampling method which is purposive have been discussed in sections 4.4.1.1, 3.4.1.2 and 3.4.1.3. Rose (2015) points out three main things that shape a sample size required in quantitative research. These points are (1) the size of variation in the value revealed by the sample and the actual population, (2) the level of the common characteristics shared by entities in the population and (3) the level of assuredness in conclusions drawn from the sample representing the population. Further details about the factors determining the sample size in the quantitative research are given in Rose (2015, pp. 198-199). In quantitative research, the sample sizes are calculated using formulas in which the application of these formulae are influenced by other factors; Rose (2015, p. 199) shows the details of these factors. However, qualitative researchers use different methods to resolve the sample size. A saturation level is reached when further analysis of the studied objects does not reveal new information or findings (Rose, 2015). The mechanism that decides the sample size in qualitative research is, “to look at what other writers have done” (Rose, 2015, p. 199). This shows, in qualitative and quantitative research, researchers use different ways to determine the sample size.

In the qualitative content analysis of the student workings, it was challenging to decide when the saturation level was reached given that similar or the same patterns of errors and solutions strategies emerged after the analysis of between 10 and 20 scripts. However, since different examination scripts have been analysed, it was hard to reach a reasonable saturation level; consequently, all the available examinations’ scripts, (for which there were 258), were explored qualitatively. Analysing this number of scripts qualitatively (1) should give the validity of conducting a statistical analysis in the quantitative phase and (2) adds strength to the acceptability of the findings in the statistical analysis of student workings.

All scripts have been considered as a sample representing the target population that also includes the future population. As for the similarity of the scripts between one semester and another, the findings in the qualitative phase have showed the occurrence of very similar patterns of data in all or the majority of the scripts; these scripts were combined from six examinations that were performed in five different semesters (two different examinations were given in one semester).

All the examinations' scripts belong to a subject which is taught in many universities internationally. These scripts are the target of the research and the inference from the analysis could be generalised to past, present and future student workings. Finally, this section completes the description of the sampling steps to start the analysis of scripts.

4.4.1.5 Implementing and reporting the sample plan.

The implementation of the sampling plan and reporting are achieved at different times during the research period. One of the requirements that shows the implementation of the sampling process is reporting it. According to Rose (2015) this "should include reporting of the target population, the sampling frame, the calculation of the sample size and response rate and details of the sampling method used" (p. 190). The implementation of the sampling method was achieved when searching for the appropriate research philosophies. The sampling unit identification will be implemented during the script collection and management step that will be covered in section 4.4.2. The establishment of the sample size will be achieved during the qualitative analysis of the scripts. The reporting of the sampling procedure is accomplished when forming and implementing steps in the sampling plan. In this research, the reporting of the sampling procedure has been detailed in sampling steps one to five that have been explained in section 4.4.1.

4.4.2 Collecting and managing raw data.

Collecting and managing the raw data (student scripts) were implemented as follows:

- At this step, collecting and managing the raw data, researchers need to familiarise themselves with the data to be analysed. Elo and Kyngäs (2008) propose that researchers should entirely explore the units to be analysed to comprehend the contents, so meaningful categories and/or concepts can be developed. The familiarisation with documents also "involves skimming, reading (thorough examination), and interpretation" (Bowen, 2009, p. 32), 2009). In this research, skimming of the scripts was performed during gathering and managing them. The details of the exploration of scripts is accomplished in the scripts' data organisation phase (section 4.4.4).

- Using the criteria defined in step three (identifying sampling unit), the scripts were scanned to check for the validity of the contents for the analysis. At the initial scanning, 500 scripts have been gathered and examined for the inclusion of suitable scripts. Table 4.1 shows that from the total number of collected scripts, only 258 scripts have passed the scripts inclusion criteria test. The reasons for the exclusion of 242 papers are shown in the table.
- The details of each script that included the examination's year, question's number, and the awarded marks to the workings were recorded and at the same time an identification number was added to each script.
- The scripts that belonged to each examination were grouped together, so there will be analysed independent of the other examinations' scripts. There are three reasons for investigating each examination's scripts separately: (1) each examination is an independent entity having its own question, (2) each script belongs to an individual student and therefore, (3) for the purpose of statistical analysis of student workings, the record of each student workings should be kept separately.
- Multiple scanning of scripts was made to check for the relevancy of the selected scripts. These further scanning had excluded the scripts that did not fulfil the script's selection criteria set down in section 4.4.1.3. At the same time the frequent scanning of scripts has also led to the modification of the initial scripts' selection criteria.

Table 4.1 Number of scripts representing student works

Years	**Number of papers analysed**	Number of works with zero mark	Number of works not attempted	Number of works excluded to reduce the analyses' load	Actual number of papers collected
2011	**27**	13	0	4	44
2011	**25**	10	3	2	40
2012	**63**	24	6	12	105
2013	**47**	17	8	21	93
2014	**32**	27	39	19	117
2015	**64**	7	15	15	101
Total number of papers	**258**	**98**	**71**	**73**	**500**

4.4.3 Preparing raw data.

To prepare the scripts for analysis, the following steps were performed.

- The pages that contain the working to be assessed were copied from each student examination's script; then the identification was added to the copied document to link it to the original student script.
- A separate data gathering document was made to collect the data from the copied document.
- Correct solutions for all questions, which students were assessed for, were prepared by the researcher.

4.4.4 Organising Analysed data.

The raw data that has been prepared for the analysis requires organising it qualitatively. The action of organising data qualitatively involves “open coding, creating categories and abstraction" (Elo & Kyngäs, 2008, p. 110). In qualitative content data analysis, open coding or simply coding refers to the action that focuses on identifying and summarising the data in the studied objects and recording the summary (Elo & Kyngäs, 2008; Rose, 2015). It is possible to form different levels of codes when coding a raw data. This hierarchical type of coding method is useful “to analyse data at varying but related levels of abstraction” (Rose, 2015, p. 364); and to gradually develop categories and subcategories (Rose, 2015). This hierarchical technique is the reverse of the abstraction process and it is helpful to conduct detailed analysis (Rose, 2015). Thus, coding, categorisation and abstraction processes may possibly stimulate researchers’ thinking resulting in forming valuable thoughts that require attention.

Researchers’ thoughts and reflections that occur during the research activities can be captured using a technique such as *memoing.* According to Rose (2015), memoing is a way of taking notes in a form of a memo which “may be a short phrase, a sentence or even a paragraph” (Rose, 2015, p. 345); seizing these notions as they arise before they disappear quickly is very important, particularly for the activities of long duration, and challenging to remember. These captured thoughts and reflections assist in driving knowledge that help answering research related problems.

The memoing technique has been extensively used due to the nature of student workings requiring thorough and rigorous analysis and interpretations. Throughout the analysis of each script, it was essential to analyse, interpret, assume or reflect why an action has occurred. New or same ideas might appear in an individual script and as the number of the analysed scripts increases the challenge of remembering all the emerging thoughts might also rise. If these reflections or thoughts were not recorded immediately, it would be very hard to recall them later; therefore, the thoughts were recorded in memos and later used, in conjunction with the data extracted from scripts to form categories and abstraction, in writing the report and forming the questionnaire.

The scheme of hierarchical coding was implemented to analyse the student scripts comprehensively, so that significant details that help understanding their workings can be developed. The following actions have been performed during the scripts analysis that involves coding scripts and recording down the produced codes:

- Compare the student workings with the researcher's solutions belonging to the original questions given to the students.
- If the solution is the same, extract the relevant data, and record it in the coding document.
- If the solution is different because of errors, for example miscopying the original question, record the error, ignore the error made and evaluate the student workings to check if the student's solution for the wrongly copied question reveal the domain knowledge for which student is being assessed. In this case, the researcher solves the modified question or part of the question; then assess the student workings by comparing the researcher's new solution with the student workings.

The analysis, using the comparison method, continued till the analysis of all the student workings were completed. Further investigation of the scripts and code recording continued to ensure the collection of all the pertinent data from scripts.

The first scanning of student workings uncovered two main patterns of data which were solution strategy (SS) and errors. The coded data were reviewed to extract new data; then the extracted data were regrouped to form categories and subcategories. Initially general categories with less details were formed. As more revision of the extracted data was performed, more subcategories and new different categories have emerged; these new entities were captured and some of the early formed categories with less meaning were removed. This revision and/or iteration processes have been applied to all the analysed scripts.

Finally, the formed categories and subcategories were accepted as the relevant form of data that represent the results of the analysed student workings. These categories and subcategories were used to understand the context that has been examined by forming an abstraction of the context. Abstraction can be made from these categories and subcategories. Using the abstracted knowledge, the nature of the analysed context can be explained. For example, the reasons that ICT tools face challenges when assessing solutions of MSQ can be explained. In the inductive content analysis of student workings, shown in Figure 4.1, the final step is reporting the result of the analysis. This will be covered in chapter five.

The understanding that was gained from the qualitative analysis of student scripts can be enhanced and/or supported further by performing the quantitative analysis of the scripts. The process and actions that achieve this aim will be described in the following sections.

4.5 Quantitative Student scripts Analysis Approach (Quantitative Methodology Designs)

The purpose of this quantitative phase is to support the qualitative phase outputs with the numerical data. In this section, numbers of quantitative research strategies will be investigated to select the most appropriate quantitative research design. Experimental and non-experimental research designs are the main strategies in quantitative research (Creswell, 2013). The researcher has explored the purposes of the two designs and selected the non-experimental design which itself contains multiple sub-designs. The main purpose of non-experimental research is to examine variables without manipulating them (Rose, 2015). In other words, “Non-experimental research designs are appropriate when the goal is to examine naturally occurring attributes, behaviours, or phenomena that cannot be experimentally manipulated by the researcher” (Laura M. O'Dwyer & James A. Bernauer, 2014, p. 152). In the quantitative process of analysing scripts, there is no manipulation of variables. Examples of these sub-designs are content analysis, natural experiments, correlational studies, and surveys (Rose, 2015). After examination of these sub-designs, the researcher has found the survey method to be more suitable than the other methods to analyse student scripts, and the rationale for selecting this design is addressed in the following section.

4.5.1 Quantitative Survey Design.

In a survey design, data is collected from a sample of a population, using a structured form of instrument such as a questionnaire, and/or a well-planned form of observation, then analysed statistically to provide the quantitative explanation of the population for the purpose of obtaining generalisable attributes about the population or drawing inferences from the results of the analysis (Creswell, 2013; Rose, 2015). According to Rose (2015) the survey study design is used in two ways: (1) descriptive survey which its purpose is to describe the phenomenon under investigation “in terms of the distribution of relevant variables within a particular population” (p. 122) and (2) analytic survey that is utilised to study the relationships between variables.

The descriptive approach includes cross-sectional and longitudinal studies; cross-sectional study covers the data that is collected at one point in time or spanning for a short period. In this research the source of data to be analysed, by survey design, comes from the qualitative data that is produced from the analysis of student scripts. The descriptions of the student scripts have been given in sections 4.3 and 4.4.2. From the descriptions, the student scripts are composed from five different semesters. This shows that the descriptive approach in the survey design is cross-sectional. The survey design using a questionnaire instrument can be constructed without requiring external efforts other than the researcher.

The general survey purposes of descriptive and analytical approaches make the survey design relevant for the analysis of scripts’ content using descriptive and inferential statistics.

According to Prasad (2008), content analysis is the questioning of the product of a kind of communication; and content analysis is considered “a method of observation in the sense that instead of asking people to respond to questions, it “takes the communications that people have produced and asks questions of communications “(Kerlinger, 1973)” (p. 2). The asking of ‘questions of communication’ might imply that either analysing the already produced data that has been created or collected or analysing previous research output for new purposes or to add more meaning to the existing research output. A transcript of interviews that require further analysis is an example of analysing collected data. In this research, the assumption that has been made is that collecting data from the student scripts, using a questionnaire instrument, would be similar to collecting data from students using interview method. This conclusion is based on the Prasad’s argument that has been made above because the student scripts content to be analysed were produced by students who had participated in the examinations. Thus, the application of the survey method to analyse student scripts could be justified as asking of questions of communication. Here the researcher questions the content of the student workings which are students’ communication of their knowledge presented in the form of scripts. The next section reveals the type of data that deductive approach can be applied to; and it also examines the instruments used to collect the data in this survey design.

4.5.2 Quantitative Deductive Approach and Data Collection Instruments.

The deductive approach starts from existing theory or concepts that are driven from previous research (Rose, 2015) and “is often used in cases where the researcher wishes to retest existing data in a new context (Catanzaro 1988)” (Elo & Kyngäs, 2008, p. 112). This may “also involve testing categories, concepts, models or hypotheses (Marshall & Rossman 1995)” (Elo & Kyngäs, 2008, p. 113). Rose identifies that deductive approach can be used in gathering and analysing data when methodologies such as qualitative, quantitative or mixed methods are employed; however, Rose explains that deductive approach is generally associated with quantitative methodology. In the deductive approach, the steps in studying research are fixed. This "is consistent with the need to operationalize concepts as measurable variables and to pre-specify appropriate data collection and analysis procedures when taking a deductive approach" (Rose, 2015, p. 85). The quantitative data collection, using deductive approach, could be achieved using different types of instrument.

The appropriate variable measurement techniques are required to collect quantitative data. Questionnaires, interviews and observations are some of the survey data instruments that are used to collect data. In the present research, the focus is extracting only the quantitative information from the qualitative data of the analysis of scripts; and the option of interviewing students is not available. The usage of the observation method to collect data from the script does not seem appropriate; because nothing is there to be observed in the scripts. The examination of the questionnaire, as the survey data collection instrument, has shown the questionnaire suitability to collect relevant data from the results of the qualitative analysis of the scripts.

4.5.3 Forming Survey Instrument - Questionnaire.

The questionnaire formation followed the strategy that was adopted from Rose (2015) and this strategy is shown in Figure 4.3. Since the questionnaire was used by the researcher

only, step five was excluded. The subsequent sections will describe the steps in the adopted strategy to form the questionnaire.

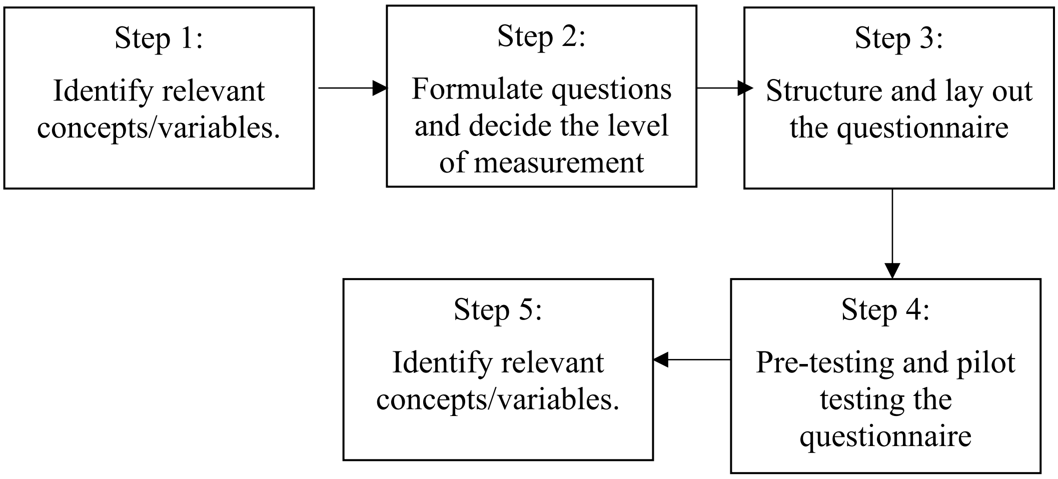


Figure 4.3 Questionnaire design process

4.5.3.1 Step 1: Identifying Relevant Concepts (Variables).

To construct a relevant research survey questionnaire, researchers are required to have adequate and appropriate knowledge. The researcher is familiar with the domain that the questions in the questionnaire investigate. Rigorous and thorough mechanisms were applied when inductive qualitative content analysis was made to analyse the student workings. The level of understanding that was obtained from the outputs of the qualitative phase was sufficient to form the relevant questions to examine student workings quantitatively; and as Rose (2015) explains that identifying relevant concepts needs paying “close attention to the definitions and dimensions of the concepts you are intending to measure” (p. 215). Therefore, the first step to identify the concepts that to be measured as variables has been achieved by evaluating the findings in the qualitative analysis and the scripts in conjunction with the research questions, the research purpose and the domain knowledge.

4.5.3.2 Step 2: Determining level of measurement and formulating questions.

The process of formulating questions in addition to locating the data sources requires establishing the level of measurement regarding quantitative variables. To enable the quantitative analysis of an observation, the observation is allocated a suitable numeric value and this procedure of giving the number to an observation is known as a measurement (Rose, 2015). Rose's statement about analysing an observation quantitatively is equally applicable to the analysis of scripts quantitatively. Therefore, the answers that are selected in response to the questions in the questionnaire can be represented in numerical form; this shows the ability to measure these answers, using their numerical values.

Madsen (2011) classifies the main data type into quantitative and qualitative. The data type that Madsen is referring to is the level of measurement that is mentioned by Rose (2015); according to Madsen, the interval and ratio data types are the subclasses of the quantitative data while nominal, ordinal and alternative are the subclasses of the qualitative data. The evaluation of the subclasses of the qualitative data revealed that the nominal level of measurement is more relevant to extract appropriate data from the scripts when compared with other types of level of measurements.

According to Rose (2015) there are two options to formulate questions: one is adopting the questionnaire that was applied in previous studies and the other is developing a questionnaire that addresses the research purpose. Riff et al. (2014) also assert that "where possible research designs may usefully take advantage of existing data gathering or variable measurement techniques that have been successfully used in past research" (p. 42). However, in the absence of a previous questionnaire, the alternative is to create a new one. Creswell (2012) states that "the substantial qualitative data collection becomes a means for developing or locating quantitative instruments" (p. 552). In this research, to analyse the scripts quantitatively, no previous questionnaire is known that could be adopted for this study. Thus, it has been necessary to develop one using the data that has been produced at the end of the qualitative phase.

4.5.3.3 Step 3: Structuring and laying out the questionnaire's questions format.

To use the identified level of measurement, an appropriate questions' layout should be selected to generate the relevant data; for nominal level of measurement, the question layout that allows the selection of one category from multiple categories is the most appropriate question format and it takes the mode of the closed question form (Rose, 2015). The advantage of the closed question form in a questionnaire is to focus the responses on the purpose of the research to obtain the specific answer. The disadvantage of using the closed question form is, it puts constraint on the respondent (Rose, 2015). However, the questionnaire in this research was used to collect answers for specific purposes, and therefore the closed question form is the appropriate type of question.

Rules are required to form clear and accurate questions. According to Rose (2015) this involves "paying careful attention to how you word your questions" (p. 224). However, the relevancy of questions is not enough because "even the best questions will not produce the desired results if your questionnaire is compromised by poor layout or a confusing structure" (p. 227). Therefore, good questions and their layouts in a questionnaire are among the features of a well-structured questionnaire. Furthermore, the information about the questionnaire's content requires management. Rose (2015) recommends keeping the track of information in the questionnaires. In the quantitative analysis of student scripts, the information about the questions in the questionnaire detailing the data they collect and the purpose of collection, the categories that describe the measured variables and the order of questions is kept in a separate document. This information was used in preparing the questionnaire and when writing the final report of the results of the analyses of the student scripts. To ensure that the data that is gathered by the questionnaire is purposeful, it is necessary to test the questionnaire, before employing it as the research instrument.

4.5.3.4 Step 4: Pre-test and pilot testing the questionnaire.

The purposes of the pre-test and pilot tests of the questionnaire are to check for its suitability to fulfil its designed purpose. The purpose of the pre-test is to ensure that the questions are well-defined and easily understood by users and/or respondents (Rose, 2015).

On the other hand, the pilot test is required to "identify any problems or difficulties with the questionnaire" (Rose, 2015, p. 231).

The questions were pre-tested to check for clarity and appropriateness. This involved the examination of the data that the questionnaire covers regarding the student scripts. Four sources were used to form questions: the domain knowledge, the student scripts, the findings of the qualitative content analysis of the scripts, and the categories and/or concepts that were developed from the findings. The questions were used to extract data from the sources. If the extracted data was available only in the raw data of the scripts and not in the findings, the findings and the categories were modified. However, if the questions did not give appropriate answers or did not adequately extract appropriate and sufficient data from the sources, the questions were modified and tested again to validate whether they collected the required data. The iteration continued until all the scripts were analysed quantitatively. The cycle of testing and modifying the questionnaire has refined the final quantitative data to be collected. The questionnaire that has been formed is in appendix A. The next task is to discuss the data that the questionnaire covers regarding the student scripts.

4.5.4 Student scripts data (questionnaire data collection).

The previous steps have led to the creation of the questionnaire. This section, briefly, discusses the data that is collected using the questionnaire. The questions cover two main subjects that are used to explain the student scripts' contents: (1) the solution strategies that are used to eliminate elements from a matrix and compute values of unknowns (variables) (the problem being solving a linear system using Gaussian elimination) and (2) the expected student errors that were made during the usage of solution strategies.

The rationale to gather the solution strategies (SS) data is to reveal students' knowledge about SSs that they use when solving questions. The SSs that were used show strategies that had been used to perform task(s) at each step. The steps might be combined by students mentally to produce specific answers. For example, step n's answer can be mentally calculated, and its answer is used in the subsequent step n+1 to produce its answer. Or the order of using SS might differ from the expected steps order when solving a question or part of a question. Furthermore, the analysis of scripts involves the investigation of types of SSs applied and the results of their application.

Students are taught to solve problems applying particular SSs. They might use correct SSs or incorrect SSs when solving problems. The correct SSs might be applied correctly to produce the required answers. However, if the correct SSs are incorrectly applied, one or multiple errors could occur. Thus, the actions of exploring and gathering the data about the applied SSs and their applications details assist in uncovering errors that are made when applying the correct types of SSs.

The incorrect application of SSs introduces one or multiple types of errors. The consequences of the occurrence of errors in the student workings introduces changes in the expected answers. Consequently, assessors interpret these errors based on their own understanding and award marks accordingly. This shows reasons for uncovering errors and their types.

The issue about the awarded marks is important when comparing student workings based on the marks they receive. The students with the highest marks make less errors or no errors at all, and sometimes the SSs they apply also differ from those awarded lower marks;

therefore, the marks are the measure of many issues in the student workings, and therefore it is important to gather data about the awarded marks. The data to be gathered is in the form of variables or categories' values that are collected as the answers to the questions in the questionnaire. Thus, it is important to establish the relationships between the research purposes, the research questions, the variables to be measured and the questions in the questionnaire. One example of the relationships between these entities are illustrated in Table 4.2. Relationships between questionnaire and quantitative research purpose

Table 4.2 Relationships between questionnaire and quantitative research purpose

Research Purpose	Research questions	Variables to be measured	Categories	Questions
To comprehend what happens in student workings	What is/are the most common errors that students make?	signError	Once, multiple	Has the signError occurred? Once = 1 Multiple = 2 If this type of error has not taken place, just leave the value of the variable blank

The next step in the process of this quantitative research process is to investigate the analysis of the quantitative data that the questionnaire gathers; and this issue is investigated in section 4.5.5.

4.5.5 Student Scripts Data Analysis.

A statistical analysis is one research method to analyse data. De Smith (2015) argues that statistical analysis is not only "purely technical exercise involving the application of specialized data collection and analysis techniques" (P. 17); De Smith relates the statistical approach to the structure of a research methodology; and this structure's characteristics determines the nature of the approach. Kothari (2004) states the role of statistics as an instrument that is used in many research activities such as "designing research, analysing its data and drawing conclusions therefrom" (Kothari, 2004, p. 131). The structure of the research methodology, in this research, applies or follows mixed methods. The statistical method is used in collecting and analysing the student workings.

The two types of statistical methods in use are descriptive and inferential. Descriptive statistics examines the quantitative information, such as population or sample data, and presents it in a summary form (Belli, 2009; Lehman et al., 2013; Rose, 2015). Inferential statistics or inferential statistical analysis is a method that is used to (1) test hypotheses, (2) "interpret the meaning of descriptive statistics" (Privitera, 2011, p. 2) and (3) exploit the sample data to draw conclusions that can be generalised to represent population's characteristics (Lehman et al., 2013; Privitera, 2011; Rose, 2015).

Descriptive statistics method will be used to analyse the student workings. This is appropriate because the main purpose of the quantitative phase is to enhance the understanding of the qualitative findings. The steps to quantify the qualitative data and form the questionnaire to collect data have been covered in section 4.5.3. In this section, the steps that will be described and implemented are shown in Figure 4.4.

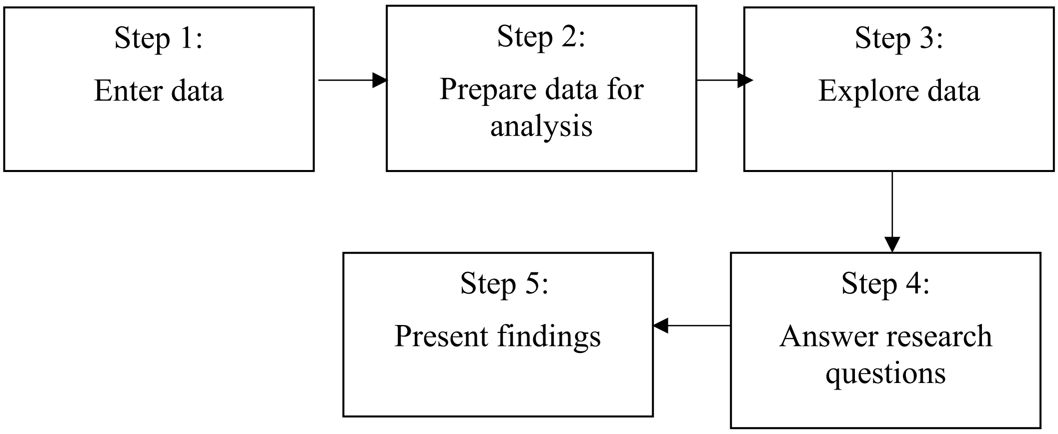


Figure 4.4 Quantitative data analysis process

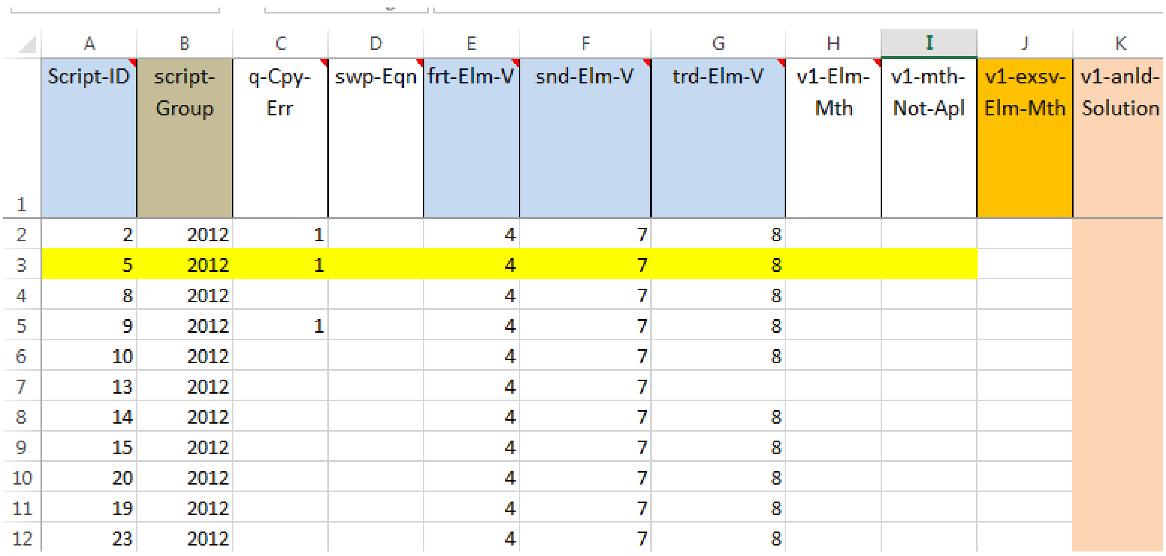
This figure shows the five steps that are performed to analyse, answer the research questions and present the findings. Steps four and five will be covered in chapter five.

4.5.5.1 Steps 1 and 2 entering and preparing student scripts data for analysis.

The data from student workings, that is to be analysed, is put in a format that permits the chosen software to analyse it. This involves the creation of a data table. In this table the first column, normally, represents a unique numerical identifier distinguishing each record that is represented by a row. The rest of the columns in the row contain the data that belongs to the records in that row. Each column represents a distinct variable. The values of variables fields come from the numerical values that are assigned to the variables' categories. In the analysis of student scripts, the coding of the variables' categories was made during the formation of the questionnaire and this table was created based on the structure and content of the questionnaire. Spreadsheets have been used to collect analysis-ready data; Table 4.3 contains an example of the coding structure while Figure 4.5 shows a section of the data that is entered into the sheets showing the implementation of the formed data table.

Table 4.3 Coding example

Question wording	Has a Double negative error taken place while executing X1 variable's elimination solution strategy?
Variable name:	v1-Dng-Mth-Apl-Err
Categories and codes	once = 1 multiple = 2 Special instruction: If no error of this type, just leave the missing value blank



	A	B	C	D	E	F	G	H	I	J	K
1	Script-ID	script-Group	q-Cpy-Err	swp-Eqn	frt-Elm-V	snd-Elm-V	trd-Elm-V	v1-Elm-Mth	v1-mth-Not-Apl	v1-exsv-Elm-Mth	v1-anld-Solution
2	2	2012	1		4	7	8				
3	5	2012	1		4	7	8				
4	8	2012			4	7	8				
5	9	2012	1		4	7	8				
6	10	2012			4	7	8				
7	13	2012			4	7					
8	14	2012			4	7	8				
9	15	2012			4	7	8				
10	20	2012			4	7	8				
11	19	2012			4	7	8				
12	23	2012			4	7	8				

Figure 4.5 Section of the collected data

The data was then imported into SPSS which was used to analyse the data. In this data table, each row represents a script representing one student workings, while the columns represent the variables that are formed from categories and subcategories. These categories and subcategories are constructed from the findings of the qualitative content analysis of the scripts. From the total of 258 scripts that have been analysed qualitatively, only 160 scripts will be analysed quantitatively. These scripts examine the same knowledge while the excluded scripts only contain partially identical knowledge to the selected ones.

Each of the student scripts that was selected for analysis contained workings which answered one multi-steps question. The full marks for these questions were different across scripts; there were two different full marks. One full mark was 5 while the other one was 6. To simplify the comparison of student awarded marks, each mark was converted to the range 0.0 -1.0.

The analysis of the awarded marks shows the number of students who were awarded different marks. Table 4.4 and Figure 4.6 shows the ranges of marks that were awarded to students and the number of students who were awarded these marks.

Table 4.4 Awarded marks and number of students

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.20	11	6.9	6.9	6.9
	.33	1	.6	.6	7.5
	.40	26	16.3	16.3	23.8
	.50	6	3.8	3.8	27.5
	.60	21	13.1	13.1	40.6
	.67	14	8.8	8.8	49.4
	.80	18	11.3	11.3	60.6
	.83	4	2.5	2.5	63.1
	.90	4	2.5	2.5	65.6
	1.00	55	34.4	34.4	100.0
	Total	160	100.0	100.0	

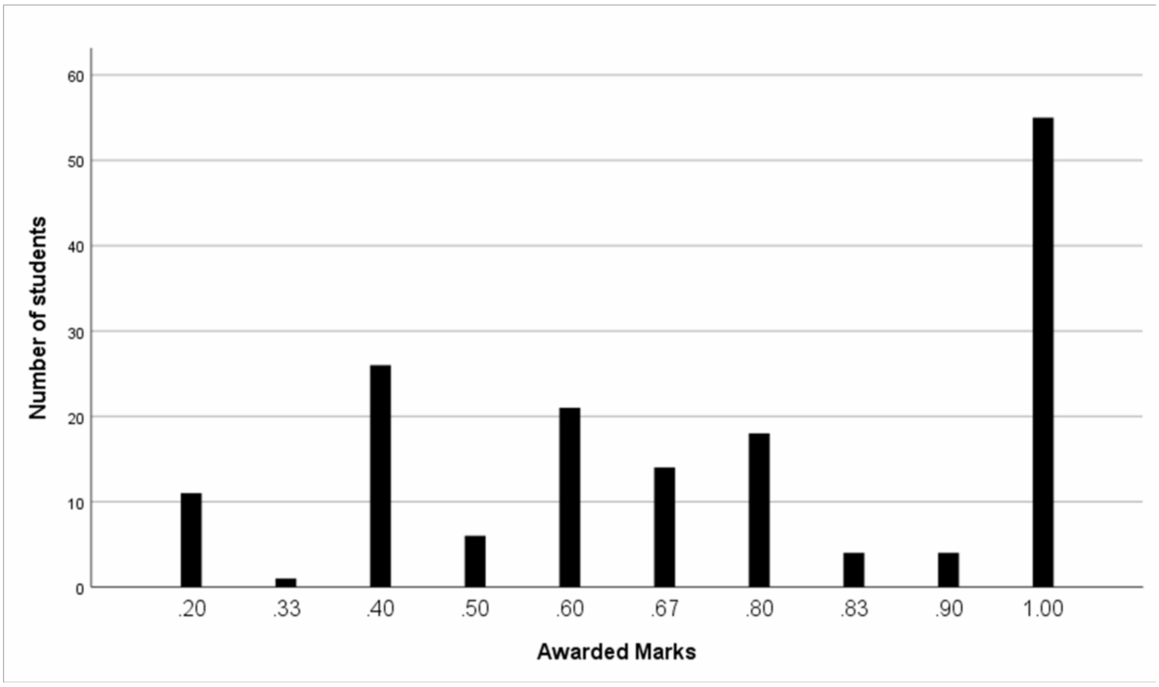


Figure 4.6 Awarded marks to students

The exploration of the analysis in the table (and graph) show that the lowest mark which is 0.2 (20%) of the total mark was awarded to 11 (2.9%) of students. The highest mark which was full mark was awarded to 55 (34.4%) students. The second highest number which is 26 students were awarded 0.4 (40%) mark. This shows the scripts that will be analysed consists of the student scripts that indicate different knowledge levels. Analysing student workings with varying data is one of the issues that assists in achieving the reliability of the process of analysing these scripts. The reliability of the process that is used analysing scripts will be discussed in section 5.1.

The analysis of the scripts data preparation is obtained by examining the data that is entered into the spreadsheets several times for input errors (including missing data) to ensure that it is correct and valid.

4.5.5.2 Step 3 Exploring student workings data.

The data exploration process involves checking the data in the table for mistakes or errors that might influence or change the expected output of the analysis of scripts. The exploration searches for: missing variables values, unexpected variables values and variables values types, for example a decimal number value instead of an integer value. The data exploration is “useful for data curation, quality control, guidance, and early intervention if applied during the data acquisition phase of a project” (Bartsch, Thompson, Jernigan, & Dale, 2014, p. 1). In this research, the exploration of the data in the table before performing the final analysis ensures that the final data to be analysed is relevant and formed correctly.

The univariate analysis which involves exploring one variable individually (Faithfull, Rodríguez, & Kuncheva, 2019; Rose, 2015) was performed on each variable’s value in the table. The result of the exploration was analysed to look for an unexpected output. One of the causes of unexpected output is the analysis of a variable value that is dependent on another related variable’s value. For example, errors and solution strategies are related- an error should not be present without student workings which requires a solution strategy – then it is expected that both such variables will have values in the table. Hence, when checking data (values) that are prepared for analysis, it is important to verify for the existence of values for related values. Thus, searching for the existence of unexpected output and correcting the value to be analysed ensures that the collected data is fit for analysis.

The questionnaire and the information that was in the memoes (whose purpose and creation were discussed in section 4.4.4) were used to correct the mistakes or errors in the table. The data exploration process used different techniques that allow combining different or related variables and analysing them based on the conditions such as variables' categories values, presence or absence of a variable or values of the variables. Finally, a frequency analysis was performed to determine the variable values, and distribution of the data variables was performed, and the analysis is presented in chapter five.

4.6 Summary

In the previous chapter, the sequential exploratory mixed methods methodology design was chosen to analyse the student scripts. This design involves the usage of both qualitative and quantitative in the same research. The purpose of this chapter was to investigate qualitative and quantitative research designs to select the most appropriate methods and use them in analysing the scripts. The investigation of the qualitative research designs produced the inductive content analysis design as being the most appropriate design to analyse the scripts qualitatively. The inductive content analysis steps that have been used to analyse the scrips have been extensively discussed. These steps included samplings steps which described issues that determine the acceptability of the scripts as a sample of a target population and the selection of purposive sampling method. The sampling steps also involved identifying a sampling unit, which identified an individual script as being the object that can be studied. In the sampling process the criteria that enabled the selection of individual script were also determined. The subject of sample size in both qualitative and quantitative phases have been discussed and implemented. In this research, the areas (sections and parts of this project) that have implemented the sampling matters were also identified.

After the description of the sampling steps the scripts were collected for analysis, which focused on searching, extracting, categorising and abstracting the scripts data. During the qualitative analysis of the scripts a memo was created and used to record the reflections of the researcher.

After the completion of the qualitative phase, the quantitative research designs were investigated, and non-experimental survey design was found to be the most appropriate approach to analyse the scripts. The discussion about the usage of the survey design was made; the review of previous publications has showed the relevancy of performing statistical analysis of the content of the scripts. This review has been required due to the general perception of the usage of the survey design to collect data only from participants or respondents, and not from a content such as documents.

The chapter also covered the deductive approach and its relationship with quantitative research. The searching for the appropriate technique to collect quantitative data resulted in selecting the questionnaire technique as being the most relevant. The absence of a previously used questionnaire that could be modified and used has led to the creation of a new questionnaire. Steps to create and use the new questionnaire were explained. The resources and mechanism to determine variables to be measured have been explained. The nominal data measurement scale was selected as the most appropriate data type and it was used to label variables. The sources of the questionnaire content and the iteration process that have been performed to form the questionnaire have been described in detail. During the creation of the questionnaire a separate memo or document was created to keep the record of the questionnaire detail and the purpose of collecting the data. The testing of the questionnaire and the general data that the questionnaire collected were also identified and the questionnaire was tested to ensure that it fulfils its purpose.

The general purpose of statistical analysis in research has been explored and the descriptive statistics was selected to analyse the student scripts. The analysis was started with the creation of a table and entering the data into the selected statistical analysis tool.

The data was explored to ensure that when analysed it would produce the expected output. In this research, it has been determined that the analysis of the data for the purpose of producing frequency distributions was sufficient to show the significance of the findings that have been produced from the qualitative analysis of the scripts. The report and the discussion of both the qualitative and quantitative analyses will be covered in chapter five.

Chapter 5 Student scripts analyses results and discussion

The previous two chapters (three and four) have investigated research process and data analyses which provide the background needed to formally investigate student workings. In the research process chapter, the needs of DSR for other knowledge to conduct its research has been established; in the same chapter, the review of the selected DSR frameworks has resulted in producing the particular DSR framework that is implemented in this research. In this framework, one of the components is the investigation of a phenomenon that helps understanding the problems and produces the knowledge that is needed to develop DSR artefacts. The philosophical assumptions to select the appropriate methodologies have also been investigated and established in the research process chapter. The data analyses chapter (chapter four) has covered the investigation of the methodologies' designs to select the relevant research methods to analyse student scripts. The analyses process that employed both qualitative content analysis and quantitative statistical analysis have been discussed in chapter four.

The qualitative and quantitative analyses of the student scripts have been explained in the previous chapter. The last phase in an inductive qualitative content analysis method is reporting and discussing the result of the analysed objects, which in this case is the student scripts in this research. This chapter presents, discusses and interprets the results of the qualitative and quantitative analyses of the scripts. Furthermore, this chapter examines the status of the knowledge that has been produced from the analyses of the scripts to produce a set of theories.

Section 5.1 will discuss various techniques that are used in showing and claiming that the outcome of the qualitative content analysis is acceptable. This acceptability can be ensured and/or enhanced by demonstrating reliability and credibility of this research. The reliability of this analysis data can be demonstrated by showing the details of the analysis process. These details might include activities such as deep investigation with the student scripts, and usage of various sampling techniques and participation of different groups in samples which are the investigated objects. This section also discusses views of different authors regarding the objectivity and/or subjectivity of validity concept regarding outputs of research.

This section also explains methods that can be used to demonstrate validity of qualitative content analysis outputs and how these methods are used in demonstrating the validity of the qualitative content analysis of the scripts. This section also discusses factors that make research activities credible and presents the activities that have been made to achieve the credibility in the analysis of student scripts.

Section 5.2 and its subsections focus on some of important matters that have been observed during the analysis of scripts. This section identifies some challenges regarding the analysis of student scripts that are linked to student actions and understanding reasons assessors use in awarding marks to the student workings. Subsections 5.2.1, 5.2.2 and 5.2.3, will cover examples that will be used to demonstrate the identified challenges.

In subsection 5.2.1, two examples present instances of difficulties in understanding student workings to analyse it. The explanations show considerable efforts that have been made to understand the intention of students, so that meaningful data can be extracted, and/or satisfactory analysis result can be achieved.

In subsection 5.2.2, challenge of knowing or determining the depth of the analysis of student workings is discussed. The amount and the type of data to be gathered from the student workings is one of the challenges in the analysis of scripts. In this subsection only one example from the student workings is used to demonstrate and explain this challenge. The analysis of student workings may also depend on the interpretation by the researcher.

The subsection 5.2.3 shows one example of challenges of classifying some findings in the analysis of student workings. The explanation and/or the interpretation is required to give reasons for associating one type of findings with the most relevant class or category.

The discussions of challenges that are made by students in their workings is followed by discussing some of assessors' marking, in section 5.2.4, when evaluating student workings.

Section 5.2.4 uses two student workings to demonstrate the inconsistencies in awarding marks to student workings. The explanations in the examples show inconsistencies that assessors have made. After the discussions of examples that have showed some of the difficulties in the analysis of scripts, the chapter progresses towards the subject of reporting the findings in the qualitative content analysis of the scripts.

Qualitative content analysis findings are presented in the form of categories and/or subcategories. Section 5.3 discusses the process of naming categories and forming abstractions. The discussion includes determining sources of names of categories, purposes and reasons for giving specific names to categories. These discussed issues are employed in forming categories and subcategories using findings in the analysis of scripts. The section is concluded by explaining a method that is used to construct abstractions. After forming abstractions, the next step is reporting the abstracted categories.

Section 5.4 describes tools and methods that will be used in reporting (presenting) findings of qualitative content analysis of student scripts as categories, subcategories and their meanings. Additionally, this section explains how the student scripts' raw data and the analysis' documents are employed in showing the analysis process and in presenting the findings and categories.

Categories, subcategories and their meanings are presented using figures and tables in section 5.4.1. The analysis discussion and interpretation are used in answering the qualitative research questions that are obtained from the understanding of student workings. After full presentation of the analysis, the focus is turned to demonstrating and explaining the analysis process that has produced the analysis. Section 5.4.2 explains the entire analysis process starting from initial analysis of selected scripts to the phase of creating a category or a subcategory.

To demonstrate this analysis process, sections 5.4.2.1, 5.4.2.2 and 5.4.2.3 present and explain the process that has been used to produce some subcategories of solution strategies and error categories. Section 5.4.2.1 explains the process that has produced the ‘ignored operation’ subcategory of error. Section 5.4.2.2 explains the process that has been used to create ‘correct expected’ and ‘incorrect with info’ subcategories of solution strategy category. The third subcategory to be explained is the ‘purposeful’ subcategory of error category.

Section 5.4.2.3 shows the analysis process that involves the creation of ‘double minus’ subcategory of error category. This section begins with the review of previous publications regarding the occurrence of ‘double minus’ subcategory of error in student workings. Again, in this section, the same analysis method and components, which have been explained in section 5.4, are used to show the creation of ‘double minus’ error. After the completion of the presentation, discussion and interpretation of qualitative content analysis scripts, the next sections present, discuss and interpret the result of the quantitative analysis of scripts.

The main purposes of section 5.5 and its subsections are to present and discuss quantitative analysis of student scripts. The format that will be used in presenting and discussing the quantitative analysis will be explained at the outset of this section. The subsections of this section will be used to present and discuss the quantitative analysis data. However, before delving into the subject, the interpretation of the ‘missing’ data, in the analysis to be presented, will be given. This understanding of the meanings of the missing data helps in when discussing the data in the analysis. The quantitative analysis data will be used to answer quantitative research questions whose answers are obtained from the outcome of the quantitative analysis of scripts

Using the format that has been given in section 5.5, section 5.5.1.1 will present and discusses the quantitative analysis regarding the order of operations when eliminating elements from a 3x3 matrix. The data in this analysis informs the number of students who have selected elements for elimination as well as the positions in the matrix of these elements. Following the presentation and discussion of data about order of operations, section 5.5.1.2 presents and discusses the analysis of types of solution strategies that have been used to eliminate elements and the number of students who have used these types of strategies. Section 5.5.1.3 also presents and discusses types of solution strategies that were used in computing values of variables and the number of students who have used each type of strategy. The advantage of uncovering types of solution strategies that have been used in the tasks of eliminating elements and computing values of variables will be discussed in the sections 5.5.1.2 and 5.5.1.3.

Section 5.5.2 explains a format used to present the quantitative analysis of all types of student errors. These analyses are presented in sections 5.5.2.1 to 5.5.2.8 covering the analyses of eight different types of errors. The discussion of the analyses of errors will be presented in section 5.5.3. These error types, their places in the student workings and their frequencies of occurrences will be discussed to identify the effects of their presences on student workings; for example, when awarding marks. The next section discusses arguments that will be used to claim the findings in the analyses of scripts as a set of theories.

Section 5.6 will present a review of a scientific theory to uncover the possibility of theorising the outcome of the analyses of student scripts. This review ends by discussing relationships between theory and practice.

The chapter concludes by presenting the chapter summary in section 5.7.

5.1 Attributes of acceptable analysis Results

This section discusses the subjects of reliability, validity and credibility of the analyses processes of scripts and the results of analyses, which determine the acceptability of the research process and its output. The designs of qualitative content analysis of scripts and of the quantitative statistical analysis of scripts have been discussed in chapter four. The qualitative content analysis design was discussed in section 4.4 and its subsections, and the quantitative scripts analysis design was discussed in section 4.5 and its subsections. Figure 5.1 shows the summary of the analyses activities that have been explained in the referred sections and subsections; furthermore, the analyses of student workings data reporting are also included. To support the acceptability of the results of the analyses, it is necessary to detail the analyses processes "so that readers have a clear understanding of how the analysis was carried out and its strengths and limitations (GAO 1996)" (Elo & Kyngäs, 2008, p. 112). The explanation of the analysis process in detail enables readers to form their own perspectives of the issues of reliability, validity and credibility concerning the activities in the research and the knowledge claims that have been made.

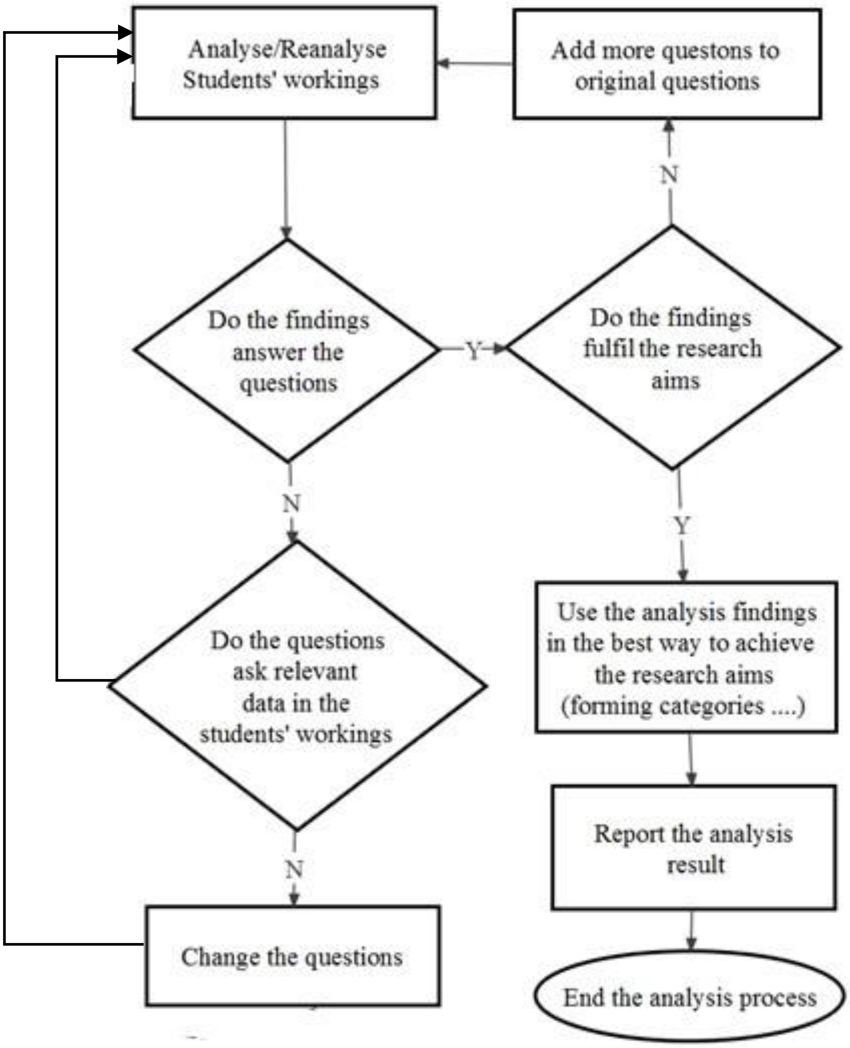


Figure 5.1 Simplified student workings analysis process

The reliability of the qualitative data, in the process of content analysis, is achieved by employing various techniques or approaches. Elo and Kyngäs (2008) identify that the reliability of the content analysis results can be achieved by showing the relationships between the results and the data that has been analysed. Elo and Kyngäs also elaborate that showing the details of the analysis process improves the reliability of the research outputs. Ebrahimi Belil, Alhani, Ebadi, and Kazemnejad (2018) further state that the factors that enhance the reliability of the output of qualitative content analyse method are: (1) to return to the inquired context (2) to make frequent and continuous contact with the inquired objects and (3) the inclusion of "maximum variation sampling technique (gender, educational status, age, employment, and type of underlying condition)" (Ebrahimi Belil et al., 2018, p. 4) in the inquired objects. The issues raised by Elo and Kyngäs (2008) and Ebrahimi Belil et al. (2018) have been achieved in analysing student workings. The relationships between the results of the analysis of student workings and the raw data of student workings can be shown as examples. Sections 5.2 and 5.4.2 include examples that show the relationships between the analysed student workings and the results of the analysis. The implementation of the subject of continuous (iterative) and persistent investigation of the inquired objects, which are student scripts, have been displayed in the activities that are presented in Figure 5.1. Student workings have been reanalysed many times during scripts collection, qualitative analysis activities, during the questionnaire formation and reporting results. For the distribution of marks and solution strategies, various student workings, containing different levels of knowledge, have been selected. The means that have been used to measure the level of the knowledge, in the student scripts, were based on the domain knowledge and the marks each student working has been awarded. Table 4.4 has shown numbers of students and marks they were given. Thus, the reliability issues have been implemented in the process of analysing scripts.

Validity cannot be separated from the reliability issue as Guba (1981) identified that "validity is a direct function of reliability" (p. 81). However, some scholars, such as Bazeley (2004), liken the validity matter to objective (quantitative) reality contradicting the necessity or importance of validity in an interpretive (qualitative) study; because as Bazeley argues the purpose of an interpretive study is for in-depth understanding. Conversely, in this research, the researcher sees both qualitative and quantitative phases playing the complementary role (see for example Seale and Silverman (1997) and explanation below). Therefore, seeking the validity of the results of analyses of scripts would not affect the in-depth understanding of the inquired objects. The act of analysing student workings, with the intention of in-depth understanding, would not nullify asking for the validity in the entire analyses processes (scripts' gathering, analysing scripts, reporting results and discussion).

The validity of the result of the analysis can be demonstrated using the same means that have been applied in establishing the reliability matter, as has been explained above. The examples that are detailed in section 5.2 can be used as one of the means to validate the output of the content analysis of inquired objects; these examples are similar to the validity of a content analysis of an object that does not require the interrater, provided that the evidence can be demonstrated as presented in section 5.2. Seale and Silverman (1997) applied this method of validation in their study and they expressed their action by saying "we do not need to appeal to interrater agreement, as the reliability and validity of the analysis is based on evidence about what the professional and client can themselves be shown to be doing" (p. 383). The professional and client are analogous to the student workings and the researcher's analysis of student workings. The other concern of validity is the link that exists between validity and credibility as Mabuza, Govender, Ogunbanjo, and Mash (2014) state: "credibility is concerned with the validity of the conclusions that are drawn from the data and how these conclusions match the reality being reported on" (p. 3). In the analysis of student workings, the credibility is established in sections 5.2 and 5.4.

The issue of the credibility of the analysis and/or research involves many aspects such as selecting scripts that have valid, reliable and credible data (student working). Credibility can be enhanced (1) by selecting appropriate inquiry objects with sufficient and acceptable data (Holmberg et al., 2018) and (2) by "identifying and describing the participants accurately" (Asplin, 2018, p. 53). These two credibility issues, concerning student workings were discussed in section 4.4.1.3. On the other hand, what weakens the credibility of research is the absence of external verification of knowledge claims that have been made (Sagitova, 2016). The means to resolve this issue of external verification has been discussed in section 3.3.3. The examples that will be given in section 5.2 show some of the implementation of what has been discussed about the credibility of the results of analyses of student scripts; furthermore, these examples show the complexity of the analysis of student workings, which has implications for achieving the research's other aim which is creating DSR artefacts.

5.2 Complexities of student workings analyses

This section reports some of the observations that have been recorded while analysing student workings. Figure 5.1 shows the iterative process that has been made until no further data can be extracted. Each iteration of the analysis action increases the validity and reliability of the data. During the iteration of the analyses of student workings, the investigations that have been made reveals the challenges of extracting all possible data and the complexity of analysing the workings.

The challenges that have been discovered are related to the analysis of student workings and awarding of marks to student workings. The complexities of analysing scripts or the challenges that have been faced during the analysis of scripts can be explained using findings in the analysis. The issues in the findings will be presented as examples and discussed; to portray the examples, the following steps will be applied:

1. Present the subject matter to be discussed.
2. Display the student script (raw data) that is linked to that subject.
3. Display the analysis' document that shows the details of the process of analysing student working to produce the analysis' result.
4. Discuss the subject in relation to what is presented by the displayed script and the document.
5. Conclude the discussion.

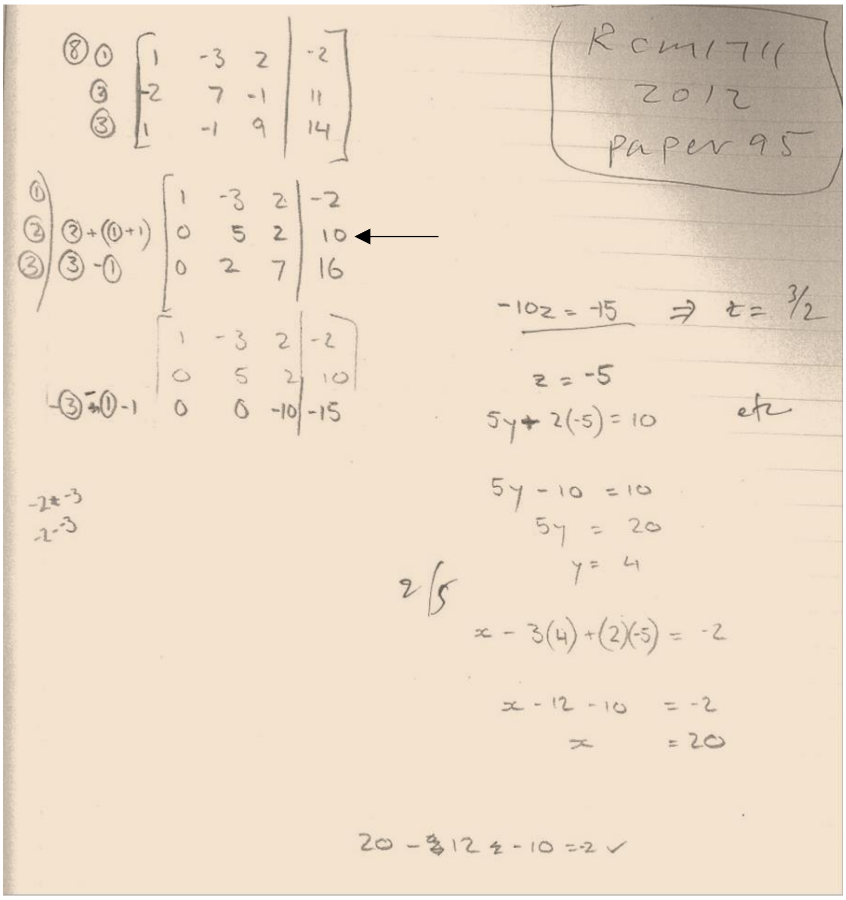
Regarding the importance of communicating knowledge clearly, Bischof and Eppler (2011) state that “the result of knowledge communication is the successful reconstruction of an insight, experience or skill by an individual because of the communicative action of another and is the more successful the more clear it is conducted” (p. 1455). Thus, the communication of the challenges, which have occurred during the qualitative analysis of student scripts, has been clarified and simplified by employing the student’s original raw data and the analysis’ document, in addition to the explanations of the subject.

5.2.1 Difficulties in understand student workings.

The analysed scripts have been selected based on the criteria that have been stated in section 4.4.1.3. However, during the analysis process, despite some scripts complying with the selection criteria, there have been many challenges in analysing them. One of these challenges is understanding student workings. The first example of this problem is shown in Figure 5.2 and Figure 5.3, for student workings, and Figure 5.4, for the analysis of the working. The explanations of the difficulties, in the example given, and how the researcher has resolved this matter are explained next.


② | ②+(①+1) | 0 5 2 | 10

Figure 5.2 Section of Figure 5.3



Rcm1711
2012
paper 95

$$\begin{array}{r}\textcircled{8}\ \textcircled{1}\\ \textcircled{3}\\ \textcircled{3}\end{array}\left[\begin{array}{ccc|c}1 & -3 & 2 & -2\\ -2 & 7 & -1 & 11\\ 1 & -1 & 9 & 14\end{array}\right]$$

$$\begin{array}{r}\textcircled{1}\\ \textcircled{2}\ \ \textcircled{2}+(\textcircled{1}+1)\\ \textcircled{3}\ \ \textcircled{3}-\textcircled{1}\end{array}\left[\begin{array}{ccc|c}1 & -3 & 2 & -2\\ 0 & 5 & 2 & 10\\ 0 & 2 & 7 & 16\end{array}\right]$$

$$\begin{array}{r}\\ \\ -\textcircled{3}-\textcircled{1}-1\end{array}\left[\begin{array}{ccc|c}1 & -3 & 2 & -2\\ 0 & 5 & 2 & 10\\ 0 & 0 & -10 & -15\end{array}\right]$$

$-2 * -3$
-2^{-3}

$\underline{-10z = -15} \Rightarrow z = 3/2$

$z = -5$ etc

$5y + 2(-5) = 10$

$5y - 10 = 10$
$5y = 20$
$y = 4$

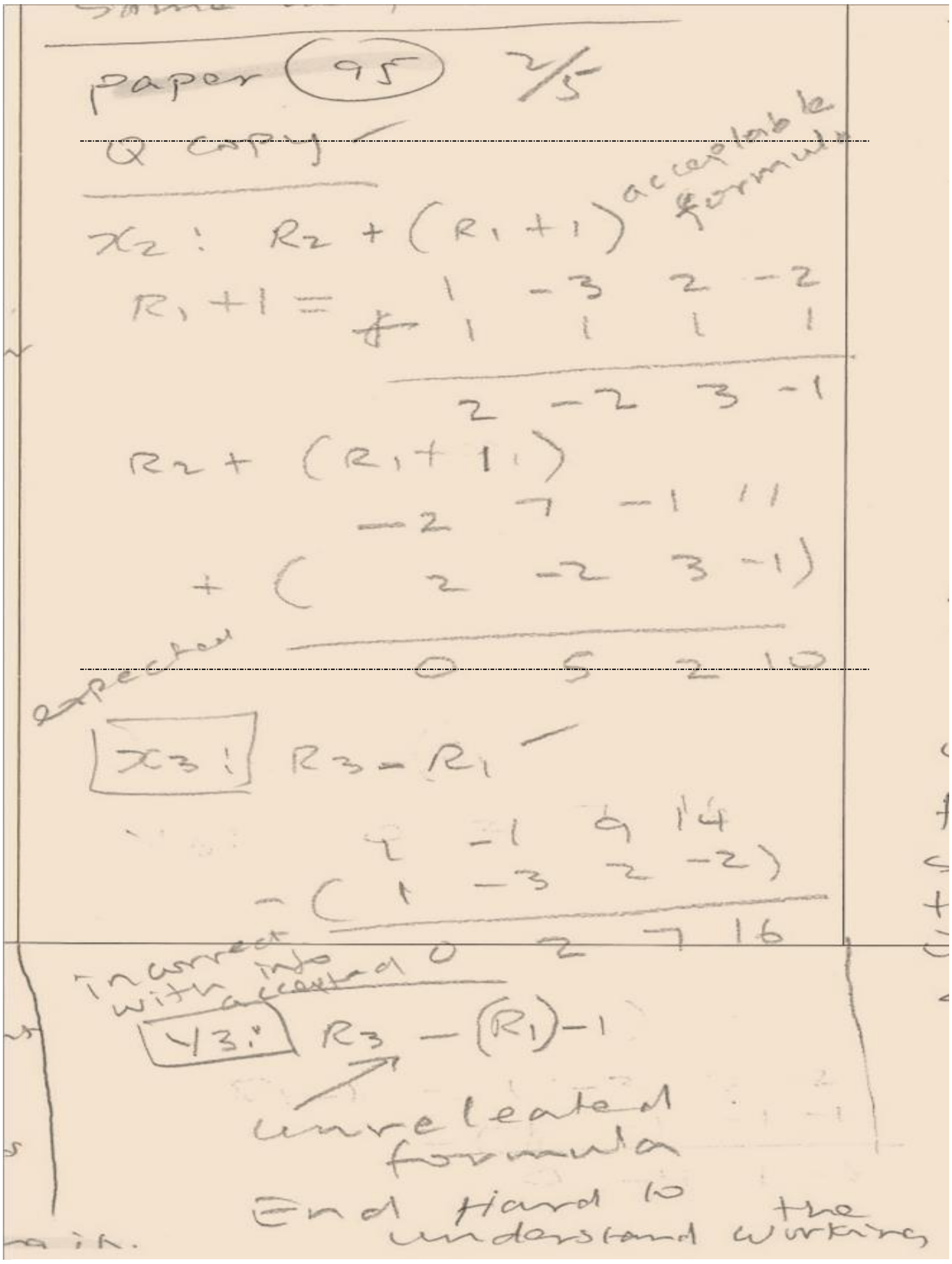
2/5

$x - 3(4) + (2)(-5) = -2$

$x - 12 - 10 = -2$
$x = 20$

$20 - 12 - 10 = -2$ ✓

Figure 5.3 Difficulty in understanding student working



paper 95 2/5

Q copy ✓

x_2 : $R_2 + (R_1 + 1)$ acceptable formula

$R_1 + 1 =$ 1 −3 2 −2
~~+~~ 1 1 1 1
2 −2 3 −1

$R_2 + (R_1 + 1)$
−2 7 −1 11
+ (2 −2 3 −1)
0 5 2 10

expected

x_3 : $R_3 - R_1$ ✓
1 −1 9 14
−(1 −3 2 −2)
0 2 7 16

Incorrect into with accepted

y_3 : $R_3 - (R_1) - 1$

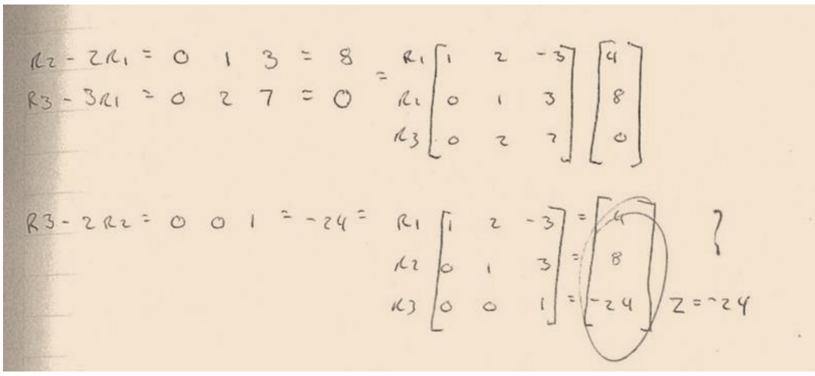
unreleated formula

End Hard to understand the working

Figure 5.4 Difficulty in understanding student working analysis' document

The focus, in this example is, on the analysis of the section of the student working that is shown in Figure 5.2; this figure has been copied from the line that has been pointed to by the arrow in Figure 5.3 that shows the complete student workings for the question that was given in the examination. Initially, the intention of the student was unclear from the initial analysis. However, further analyses identified that the intention of the solution strategy could be written as R2 + (R1 + 1). The strategy is translated into add the numerical value 1 to the equation 1 labelled as the R1 (row 1); then add the total to the R2's (row 2's) columns' values. The question is whether to add the numerical value 1 to all the columns' values in row 1 or to only one or some of the selected columns. The conclusion of the analysis was that the numerical value 1 should be added to all the elements in that row; and this assumption was transformed into the intended working, as shown on the analysis document Figure 5.4, that R1 + 1 = 2, -2, 3, -1. Then this value is added to R2's columns' values, as shown R2 + (R1 + 1) = 0 5 2 10. The two operations are shown between the two broken lines, drawn horizontally, on the analysis document. The answer that is displayed by R2 + (R1 + 1) operation matches the student's answer that has been displayed in Figure 5.2. Hence, this student script examination demonstrates the challenges that exist in the analysis of their workings and also shows the depth of the analysis required; furthermore, the script indicates that the result of the analysis of this type of content is far from being only the summary of the content of the analysed object, rather it involves the interpretation of the content.

The second example in the difficulty of understanding student workings is the 'unexplainable error'. These are student errors for which it is challenging to extract any information. Figure 5.5 is the section that is taken from Figure 5.6 which the student workings is to solve the given question. There are three different operations that are shown in Figure 5.6; these operations are performed using R2-2R1, R3-3R1 and R3-2R2 solution strategies.



$R_2 - 2R_1 = 0 \quad 1 \quad 3 = 8$

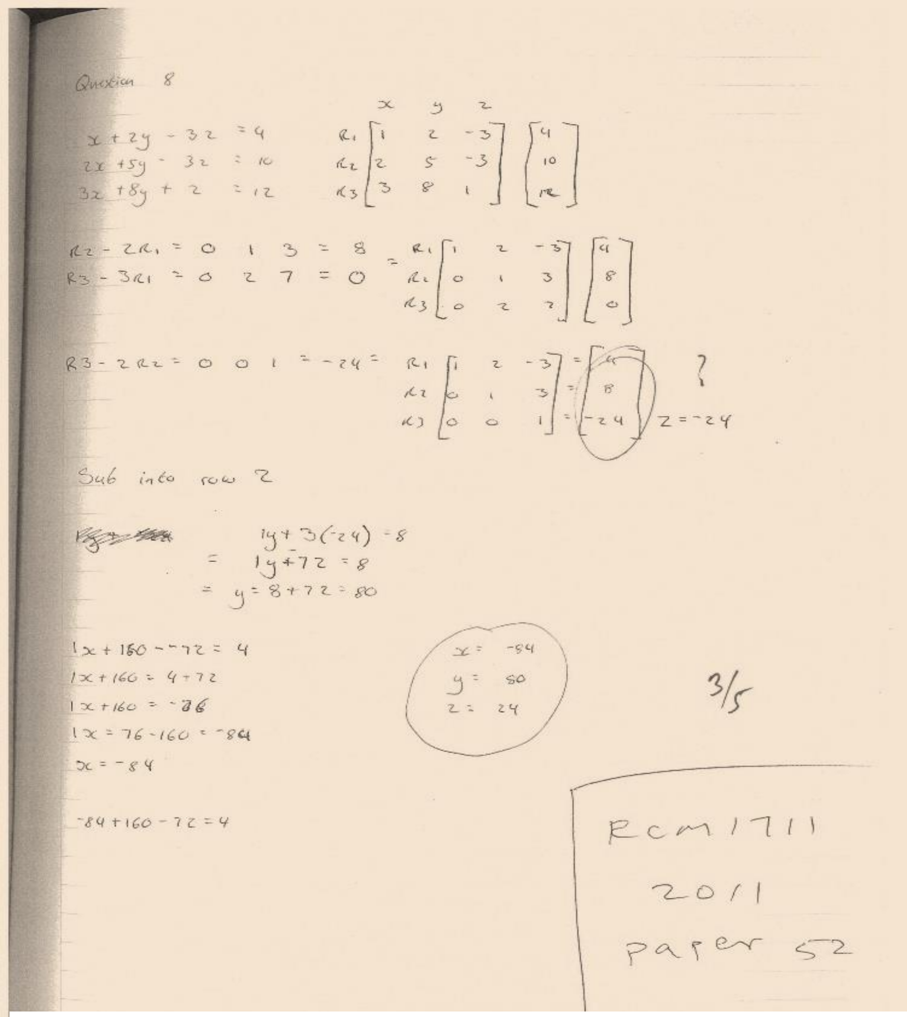
$R_3 - 3R_1 = 0 \quad 2 \quad 7 = 0$

$$= \begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 3 \\ 0 & 2 & 7 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 0 \end{bmatrix}$$

$$R_3 - 2R_2 = 0 \quad 0 \quad 1 = -24 = \begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ -24 \end{bmatrix} \; ? \quad Z = -24$$

Figure 5.5 Section of Figure 5.6

The analysis of these three operations are shown on the analysis' document that is displayed in Figure 5.7; each operation's process starts from where the solution strategy is underlined, using broken lines.



Question 8

$$\begin{array}{l} x + 2y - 3z = 4 \\ 2x + 5y - 3z = 10 \\ 3x + 8y + z = 12 \end{array} \qquad \begin{array}{c|ccc} & x & y & z \\ R_1 & 1 & 2 & -3 \\ R_2 & 2 & 5 & -3 \\ R_3 & 3 & 8 & 1 \end{array} \; \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$R_2 - 2R_1 = 0 \quad 1 \quad 3 = 8$

$R_3 - 3R_1 = 0 \quad 2 \quad 7 = 0$

$$= \begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 3 \\ 0 & 2 & 7 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 0 \end{bmatrix}$$

$$R_3 - 2R_2 = 0 \quad 0 \quad 1 = -24 = \begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ -24 \end{bmatrix} \; ? \quad Z = -24$$

Sub into row 2

$1y + 3(-24) = 8$

$= 1y + 72 = 8$

$= y = 8 + 72 = 80$

$1x + 160 - -72 = 4$

$1x + 160 = 4 + 72$

$1x + 160 = -76$

$1x = 76 - 160 = -84$

$x = -84$

$-84 + 160 - 72 = 4$

$x = -84$

$y = 80$

$z = 24$

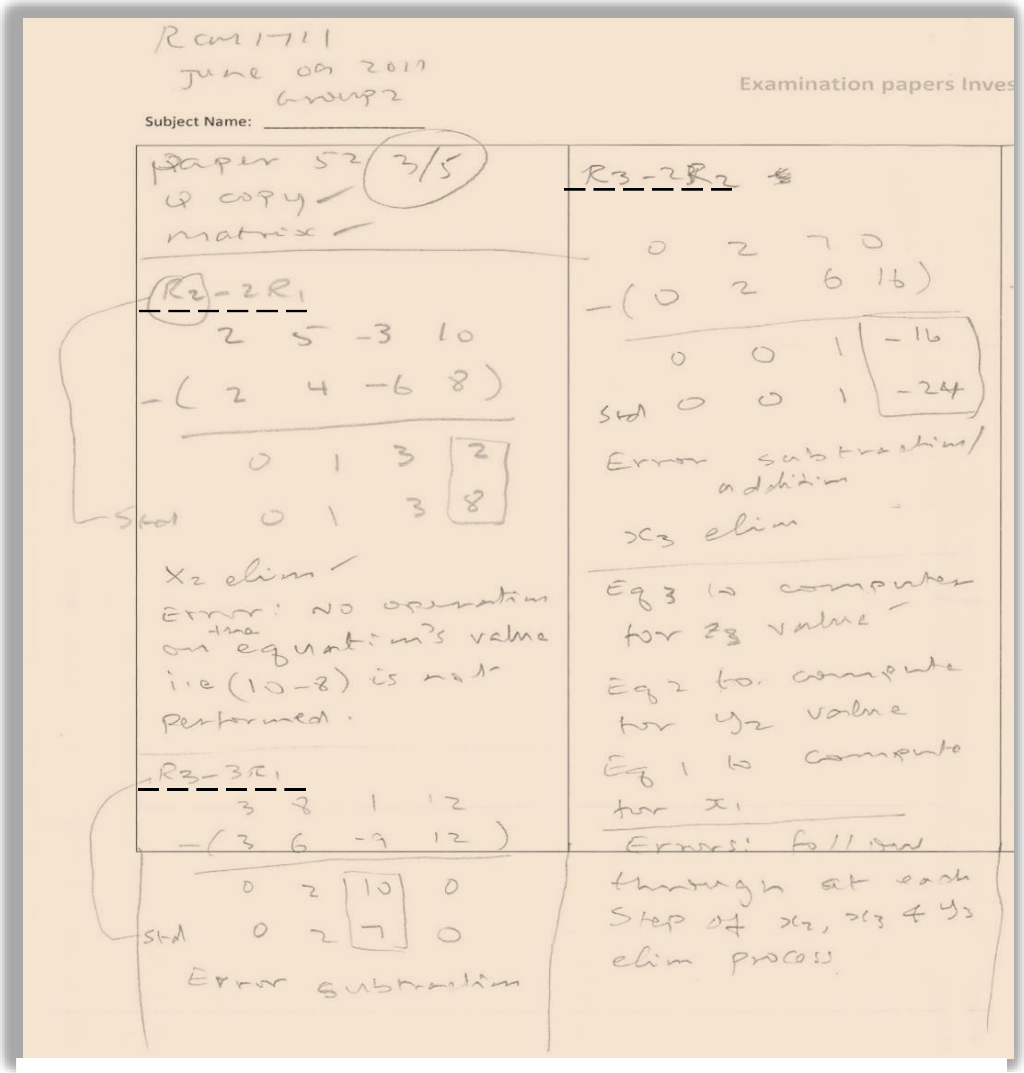
3/5

RCM1711

2011

paper 52

Figure 5.6 Unexplainable error in student working



Rcm1711
June on 2017
Group2

Examination papers Inves

Subject Name: ____________

paper 52 3/5
Q copy ✓
matrix ✓

R2-2R1
2 5 -3 10
-(2 4 -6 8)
0 1 3 2
Std 0 1 3 8

X2 elim ✓
Error: No operation on the equation's value i.e (10-8) is not performed.

R3-3R1
3 8 1 12
-(3 6 -9 12)
0 2 10 0
Std 0 2 7 0

Error subtraction

R3-2R2
0 2 7 0
-(0 2 6 16)
0 0 1 -16
Std 0 0 1 -24

Error subtraction/addition

x3 elim

Eq 3 to compute for z3 value ✓
Eq 2 to compute for y2 value
Eq 1 to compute for x1

Errors: follow through at each step of x2, x3 & y3 elim process

Figure 5.7 Figure 5.7Analysis of unexplainable error in student working

From the analysis results, each of the three operations did not produce the correct result. In the student working the R2-2R1 = 0, 1, 3, 8. The correct answer on the analysis' document is 0, 1, 3, 2. The comparison of the two answers shows that for the final value, which is the result of the subtraction 10-8, the student answer is 8 while the correct answer should be 2. Similar incorrect actions had also been taken by the student in the R3-3R1 operation; the 3, 8, 1, 2 – (3, 6, -9, 12) = should give the correct answer which is 0, 2, 10, 0. However, the student answer is 0, 2, 7, 0. The error is in the 1 – (-9) operation part. The total should be 10 not 7. The third operation is 0, 2, 7, 10 – (0 2 6 16). The correct answer is 0, 0, 1, -16. However, the student answer is 0, 0, 1, 24. The student has made errors in all the three operations, which are all subtraction operations showing something is going wrong. It is hard to give a reason; to claim that the student problem is performing subtraction would seem unjustifiable claim. Note that the subtraction errors were only made in one column's operation out of four columns' operations. This indicates that the problem does not lie with the student's ability to perform simple arithmetic. In three equations, there are subtraction operations. The student has made errors in three different raw operations. Therefore, this type of error will be classified as 'other', due to the difficulty of resolving the issue of categorising these errors.

5.2.2 Depth of student workings analyses.

One of the complexities of the analyses of student workings is the issue of the depth of the analyses, which is how much data to gather and what types of data to gather. Some student workings contain irrelevant further working which may itself be incorrect. This section is focusing on the Section of working that is copied from Figure 5.8 and is shown in Figure 5.9. The analysis for this student working is also displayed in Figure 5.10.

As shown in Figure 5.9, the student has used the R1-2R3 solution strategy to eliminate the z term from equation 3, which is not necessary. Two problems are caused in applying this strategy. First, the step to eliminate the z term is not necessary; because Gaussian elimination aims to involve an upper triangular matrix in which the z term would be ordinary non-zero. Second, the strategy is incorrectly applied. The section, which sits between the two horizontally drawn broken lines, on the analysis' document in Figure 5.10, identifies the correct result of the application of the strategy. The student's result for the application of this strategy is displayed on the last row in Figure 5.9; it is equal to 0, 0, 0, -6, while the correct answer on the analysis' document is 1, -3, 0, - 6. The analysis of the result on the document indicates the strategy has in fact reintroduced the x and y terms to equation 3 nullifying the actions of the previous two strategies that have been used to eliminate them. Thus, the consequence of the student action could be considered as he/she is changing his/her mind; therefore, this excessive action forces the assessor to either give no mark to the working that is affected by this action or give partial mark (the meaning explained in section 2.4.1) to the working for displaying the knowledge of eliminating terms from the equation.

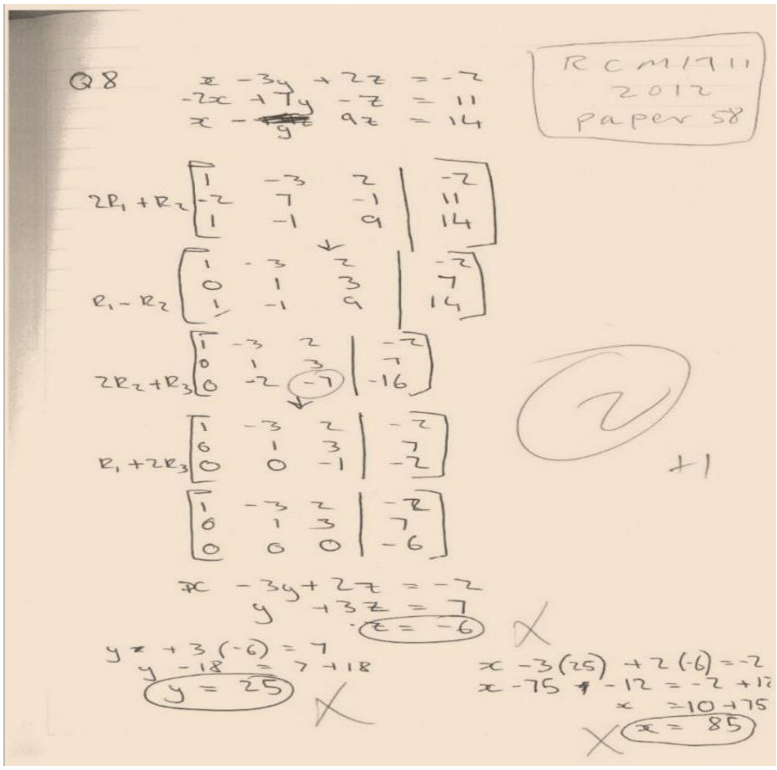
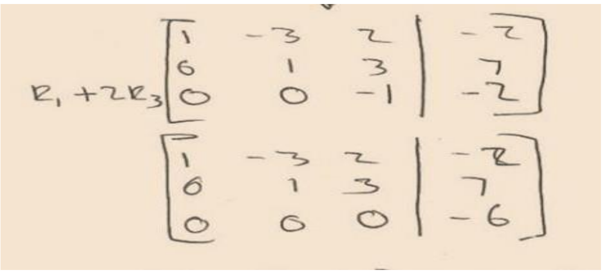


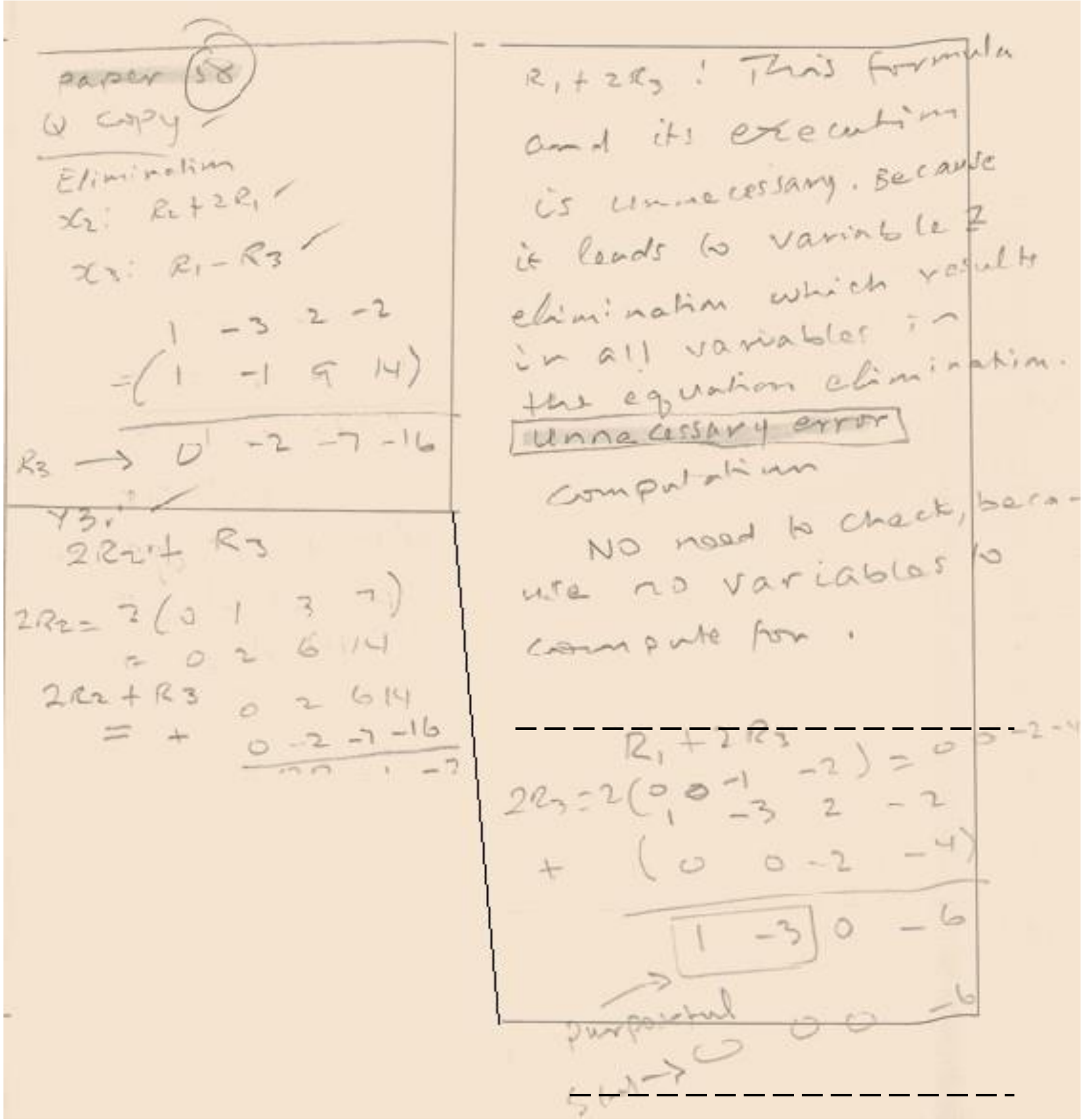
Figure 5.8 Student working showing excessive work



$$R_1+2R_3\left[\begin{array}{ccc|c} 1 & -3 & 2 & -2 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & -1 & -2 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & -3 & 2 & -2 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 0 & -6 \end{array}\right]$$

Figure 5.9 Section from Figure 5.8



paper 58

Q copy ✓

Elimination

x_2: R_2+2R_1 ✓

x_3: R_1-R_3 ✓

1 -3 2 -2

-(1 -1 9 14)

$R_3 \to$ 0 -2 -7 -16

Y_3: $2R_2+R_3$

$2R_2 = 2(0\ 1\ 3\ 7)$

$= 0\ 2\ 6\ 14$

$2R_2+R_3$ 0 2 6 14

= + 0 -2 -7 -16

R_1+2R_3 : This formula and its execution is unnecessary. Because it leads to variable Z elimination which results in all variables in the equation elimination.

Unnecessary error computation

No need to check, because no variables to compute for.

R_1+2R_3

$2R_3 = 2(0\ 0\ -1\ -2) = 0\ 0\ -2\ -4$

1 -3 2 -2

+ (0 0 -2 -4)

1 -3 0 -6

purposeful

0 0 -6

Sol → 0

Figure 5.10 Analysis of student working showing excessive work

This type of action poses the question of the depth of the analysis of student workings as well as the depth of the interpretation of the findings in the analysis. The examination of the analysis' document shows the efforts that have been made to analyse this working to deduce data to build extensive and reliable knowledge that helps understanding student workings to achieve the purpose of this research.

5.2.3 Challenges of classifying data in the results.

After producing data from the analysis, categorisation of this data is the challenging issue. For example, consider the student analysis shown in Figure 5.11, the section to be reviewed, which comes from complete student workings revealed in Figure 5.12. The findings of the analysis for the entire student workings is in the analysis' document that is given in Figure 5.13.

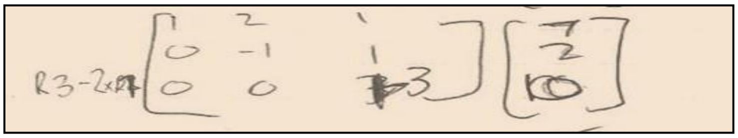
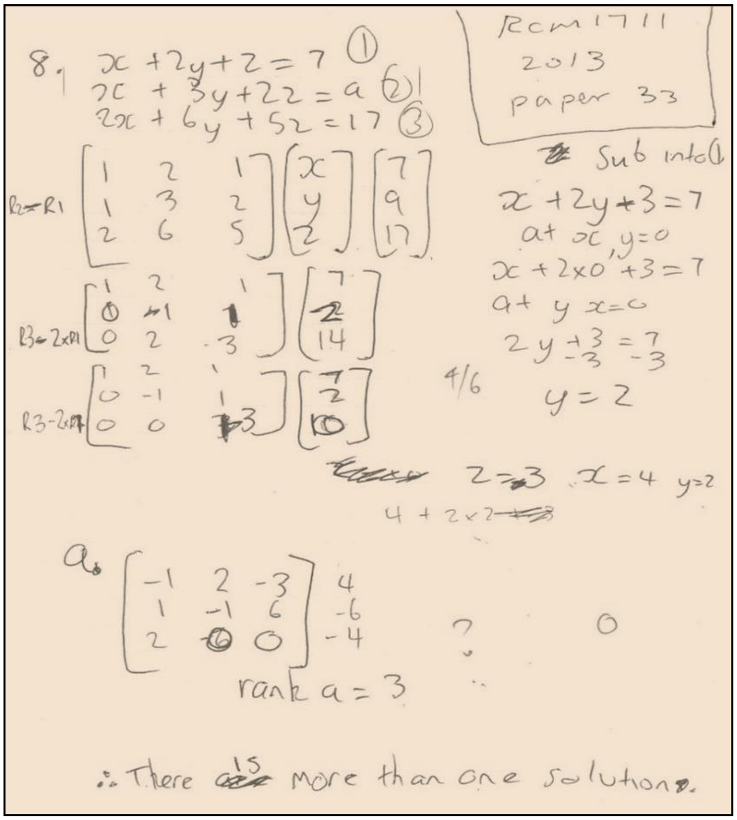


Figure 5.11 Section of Figure 5.12



Rcm1711
2013
paper 33

8. $x+2y+z=7$ ①
$x+3y+2z=a$ ②
$2x+6y+5z=17$ ③

R2-R1 $\begin{bmatrix}1&2&1\\1&3&2\\2&6&5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}\begin{bmatrix}7\\9\\17\end{bmatrix}$

R3-2xR1 $\begin{bmatrix}1&2&1\\0&-1&1\\0&2&3\end{bmatrix}\begin{bmatrix}7\\2\\14\end{bmatrix}$

R3-2xR[illegible] $\begin{bmatrix}1&2&1\\0&-1&1\\0&0&3\end{bmatrix}\begin{bmatrix}7\\2\\10\end{bmatrix}$

Sub into ①
$x+2y+3=7$
at x, $y=0$
$x+2\times0+3=7$
at y $x=0$
$2y+3=7$
-3 -3
$y=2$

4/6

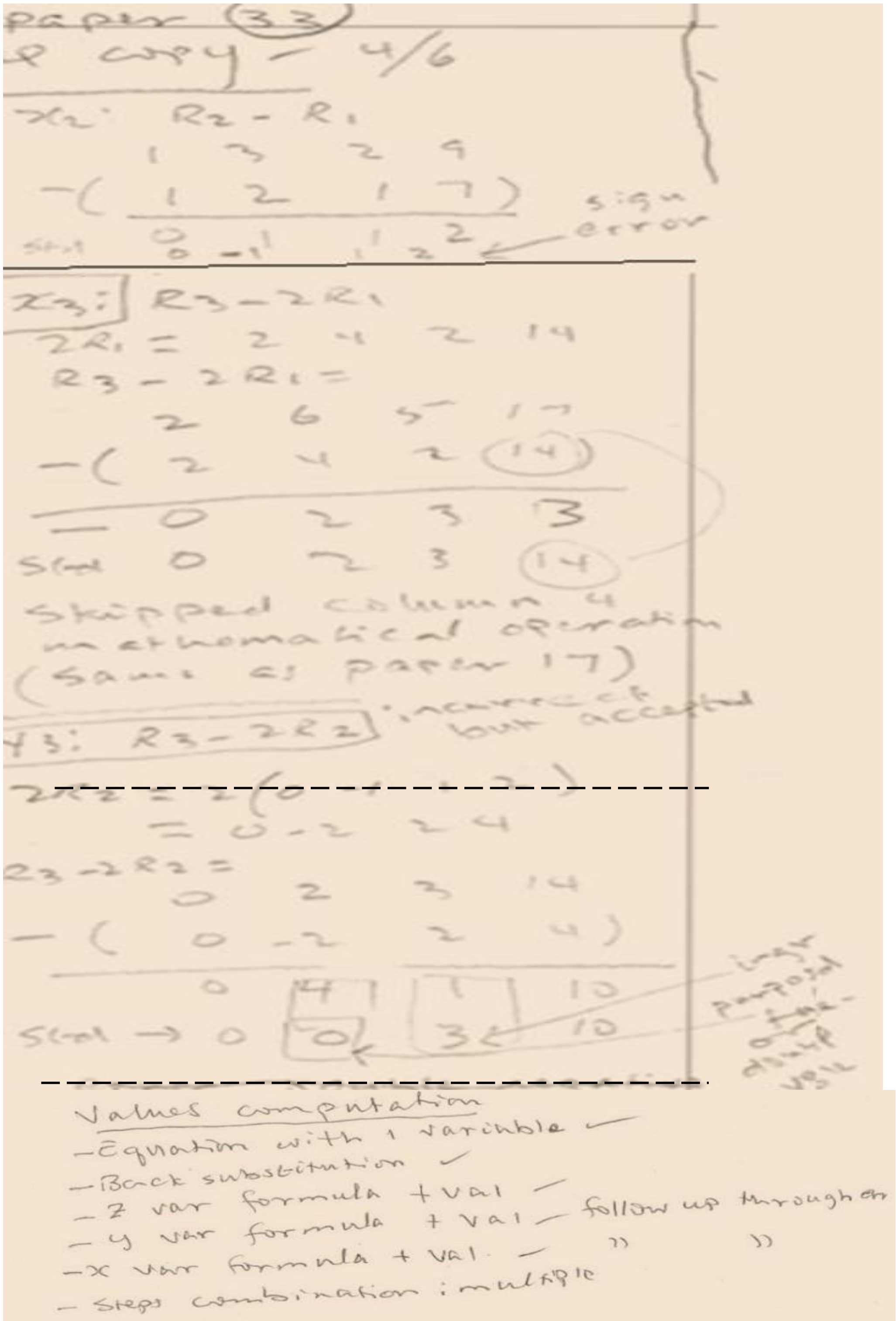
$z=3$ $x=4$ $y=2$
$4+2\times2$ ~~+3~~

a. $\begin{bmatrix}-1&2&-3\\1&-1&6\\2&6&0\end{bmatrix}\begin{matrix}4\\-6\\-4\end{matrix}$? 0

rank a = 3

∴ There ~~are~~ is more than one solution.

Figure 5.12 Student workings-categorisation challenge 1



paper (33)

copy – 4/6

x_2: $R_2 - R_1$

1 3 2 9

−(1 2 1 7)

0 1 1 2

Stu 0 −1 1 2

sign error

x_3: $R_3 - 2R_1$

$2R_1$ = 2 4 2 14

$R_3 - 2R_1$ =

2 6 5 17

−(2 4 2 14)

0 2 3 3

Stu 0 2 3 14

skipped column 4 mathematical operation (same as paper 17)

y_3: $R_3 - 2R_2$: incorrect but accepted

$2R_2$ = 2(0 1 1 2)

= 0 −2 2 4

$R_3 - 2R_2$ =

0 2 3 14

−(0 −2 2 4)

0 4 1 10

Stu → 0 0 3 10

incor proposal [illegible]

Values computation

- Equation with 1 varinble ✓
- Back substitution ✓
- z var formula + val –
- y var formula + val – follow up through an
- x var formula + val – " "
- steps combination : multiple

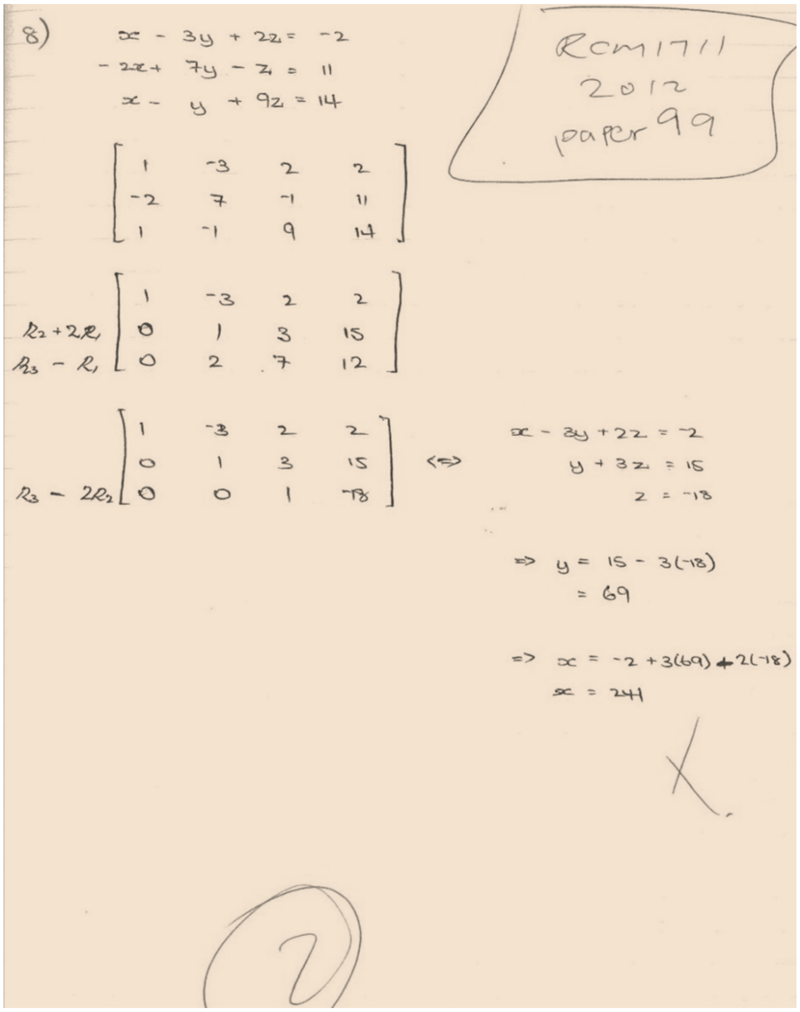
Figure 5.13 Student working analysis findings categorisation challenge 1

The analysis’ document has found three different errors with some of them requiring interpretation and making assumptions to deduce relevant meanings. However, in this section the concentration is only on one of the errors that has been encompassed in Figure 5.11. The other errors that have been identified are described in the comments on the document. The part that has been addressed in this section is concerned with the R3-2R2 solution strategy application details that is shown between the horizontal broken lines on the document.

Figure 5.11 shows the student's changing the 2R2 to 2R1 by crossing the number 2. However, the student appears to intend to compute R3 – 2R2 instead of R3 – 2R1 from the results in that row. The purpose of the strategy is to eliminate the y term from equation 3. The student output for the application of the strategy is 0, 0, 3, 10 but the correct result that is shown in the document is 0, 4, 1, 10. The cause of the error in the student's result is either from the negligence of the double negative sign, subtracting a negative number in the operation, or setting the result, purposely, to 0 to indicate the elimination of the variable. Therefore, it is hard to determine which reason is the main cause of the error. In this operation, the error is recorded in the purposeful error subcategory, since the purpose of the operation is to eliminate the variable and the occurrence of this phenomenon is widespread through scripts across different student workings groups that have been analysed in this research.

5.2.4 Inconsistency in awarding marks to student workings.

Some inconsistencies in awarding marks to student workings have been observed in the assessors' actions. There were many examples that could be shown here, however the workings that are shown in Figure 5.14, and Figure 5.15 as examples are sufficient to demonstrate this problem. The two workings in the figures have been awarded the same mark despite showing different levels of understandings. The working in Figure 5.14 reveals all the correct working that warrant receiving the full marks, but the working in Figure 5.15 did not show all the working.



8) $x - 3y + 2z = -2$

$-2x + 7y - z = 11$

$x - y + 9z = 14$

Rcm1711
2012
paper 99

$$\begin{bmatrix} 1 & -3 & 2 & 2 \\ -2 & 7 & -1 & 11 \\ 1 & -1 & 9 & 14 \end{bmatrix}$$

$$\begin{array}{r} \\ R_2 + 2R_1 \\ R_3 - R_1 \end{array} \begin{bmatrix} 1 & -3 & 2 & 2 \\ 0 & 1 & 3 & 15 \\ 0 & 2 & 7 & 12 \end{bmatrix}$$

$$\begin{array}{r} \\ \\ R_3 - 2R_2 \end{array} \begin{bmatrix} 1 & -3 & 2 & 2 \\ 0 & 1 & 3 & 15 \\ 0 & 0 & 1 & -18 \end{bmatrix} \Leftrightarrow \begin{array}{r} x - 3y + 2z = -2 \\ y + 3z = 15 \\ z = -18 \end{array}$$

$\Rightarrow y = 15 - 3(-18)$

$= 69$

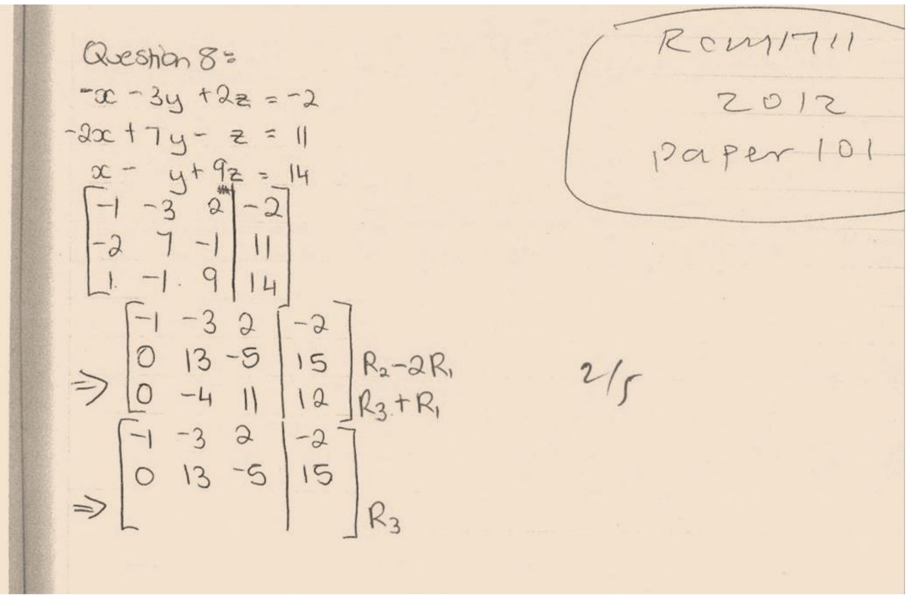
$\Rightarrow x = -2 + 3(69) + 2(-18)$

$x = 241$

X

②

Figure 5.14 Inconsistency in marking student working –paper 99



Question 8:

$-x - 3y + 2z = -2$

$-2x + 7y - z = 11$

$x - y + 9z = 14$

$$\left[\begin{array}{ccc|c} -1 & -3 & 2 & -2 \\ -2 & 7 & -1 & 11 \\ 1 & -1 & 9 & 14 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{ccc|c} -1 & -3 & 2 & -2 \\ 0 & 13 & -5 & 15 \\ 0 & -4 & 11 & 12 \end{array}\right] \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 + R_1 \end{array}$$

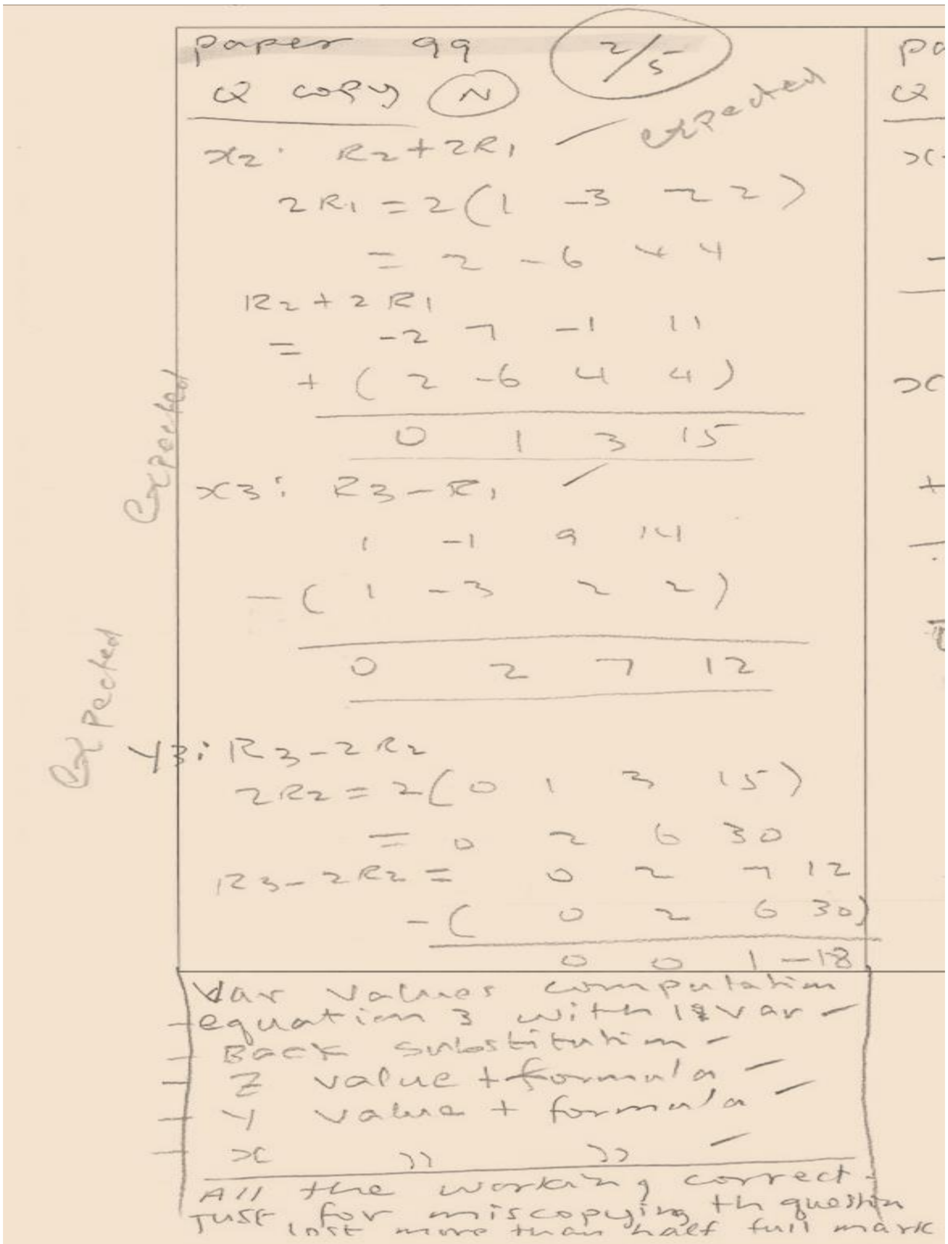
$$\Rightarrow \left[\begin{array}{ccc|c} -1 & -3 & 2 & -2 \\ 0 & 13 & -5 & 15 \\ & & & \end{array}\right] R_3$$

RCM1711
2012
paper 101

2/5

Figure 5.15 inconsistency in marking student working-paper 101

The analysed workings for the papers are shown in Figure 5.16 and Figure 5.17 respectively. The researcher has concluded that the reason for the working in Figure 5.14 losing the full marks was for miscopying the original question. However, the working shown in Figure 5.15 has lost marks for not only miscopying the original question but also for not showing all working.



paper 99 2/5

Q copy N

x_2: R_2+2R_1 ✓ expected

$2R_1 = 2(1 \quad -3 \quad 2 \quad 2)$

$= 2 \quad -6 \quad 4 \quad 4$

R_2+2R_1

$= \quad -2 \quad 7 \quad -1 \quad 11$

$+ (2 \quad -6 \quad 4 \quad 4)$

$0 \quad 1 \quad 3 \quad 15$

Expected

x_3: R_3-R_1 ✓

$1 \quad -1 \quad 9 \quad 14$

$-(1 \quad -3 \quad 2 \quad 2)$

$0 \quad 2 \quad 7 \quad 12$

Expected

Y_3: R_3-2R_2

$2R_2 = 2(0 \quad 1 \quad 3 \quad 15)$

$= 0 \quad 2 \quad 6 \quad 30$

$R_3-2R_2 = \quad 0 \quad 2 \quad 7 \quad 12$

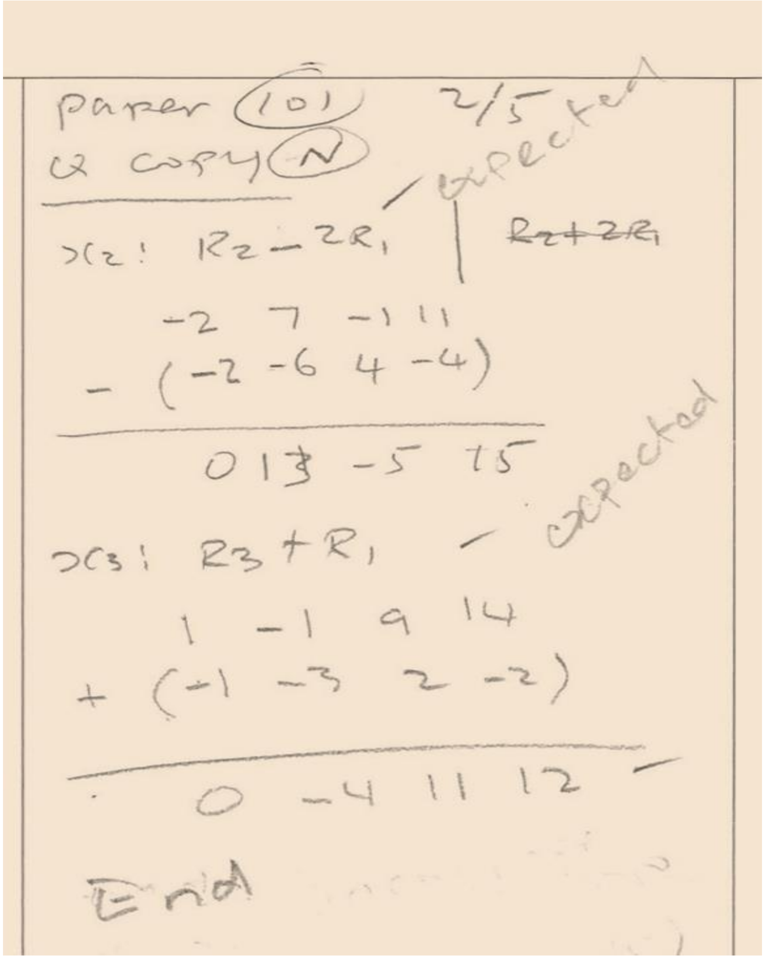
$-(0 \quad 2 \quad 6 \quad 30)$

$0 \quad 0 \quad 1 \quad -18$

Var values computation
equation 3 with 1 var ✓
Back substitution ✓
Z value + ~~formula~~ ✓
Y value + formula ✓
x " " ✓

All the working correct.
Just for miscopying th question lost more than half full mark

Figure 5.16 Analysis of student working of Figure 5.14



paper (101) 2/5

Q copy (N)

x2: R2 − 2R1 / expected | ~~R2+2R1~~

−2 7 −1 11

− (−2 −6 4 −4)

0 13 −5 15

x3: R3 + R1 / expected

1 −1 9 14

\+ (−1 −3 2 −2)

0 −4 11 12

End

Figure 5.17 Analysis of student working of Figure 5.15

Section 5.2 and its subsections have discussed some of the observations that were extracted from memo concerning the subjects that the researcher had discerned during the analysis of the scripts and the creation of the questionnaire. The following section focuses on presenting the findings of the qualitative content analysis of student workings.

5.3 Categories and subcategories formation

The final step, which has been shown in Figure 4.1, in the inductive qualitative content analysis process is reporting the result of the analysis. According to Elo and Kyngäs (2008), the result of a qualitative content analysis is the summary of the original information that is presented in the form of concepts and/or categories that express the investigated phenomenon. Categories are classes or groups of entities formed “according to their similarities and differences” (Rose, 2015, p. 99); and the data in the entities are the initial observation and description of the phenomena (Hevner & Chatterjee, 2010). In the qualitative content analysis, the source of categories are codes "that are related to each other through their content or context" (Erlingsson & Brysiewicz, 2017, p. 94). In the qualitative analysis of student workings, the findings are the solution strategies and the student errors that are made during the application of these strategies. Therefore, the data that is deduced from the analysis of the workings is not only considered the summary of their workings, as Elo and Kyngäs claimed, but in the analysis of the workings. That data involves the understandings and interpretations of the workings, which contain hidden data. The processes of applying different levels of analyses to the student workings are to form categories and subcategories that have been explained in section 4.4.4.

The names of categories/subcategories, as the output of the qualitative content analysis of a context, come from previous publications (Vaismoradi, Jones, Turunen, & Snelgrove, 2016). These names features should be "factual and short" (Erlingsson & Brysiewicz, 2017, p. 94). The names of categories, subcategories, and of concepts arising from the scripts' analysis are driven from the data that were uncovered from the investigation of student workings and the interpretation of the data during the investigation. However, new terminologies have been required due to the absence of the terms in the domain from which the student workings originated. In other words, these new terminologies are not derived from the language of mathematics which is the domain. For example, the term 'purposeful error' is not in mathematics terminology. These new terminologies are used in naming some of the subcategories of the solution strategies, and errors.

The student workings are based on the domain knowledge that examinations' questions require students to display. In answering questions, students might use specific steps or solution strategies such as 2R2-R1 to solve a task in a question. If this solution strategy were correct and applied correctly, the expected outcome (the correct solution) would be achieved. Another student might use 3R2-2R3, which is another acceptable solution strategy to perform the same task to produce the acceptable answer. The third student might use an incorrect solution strategy to perform the same task. If each possible correct solution strategy, which is used by students, is to be collected, there will be multiple different specific solution strategies to be collected for analysis. However, if a few common terms were used to group these solutions strategies, it would be easier to gather data about and categorise them under these common terms. Thus, several terms have been devised so that the gathered data can be grouped under them: ‘expected, ‘acceptable’ and ‘incorrect’ have been chosen. These will be categories’ names for grouping of the findings that have resulted from the analysis of the workings, which employed solution strategies.

The ‘expected’ category is the solution strategy that students are taught to use when solving problems to produce a valid output; the ‘acceptable’ solution strategy category also produces a correct answer; however, it contains more work than the expected solution strategy category. If the expected or acceptable solution strategy category is applied correctly, the expected product is the expected or acceptable answer. However, if the strategy is incorrectly applied, the expected output is incorrect answer, and the working contains error(s). The category and subcategories names for the errors have also been developed. The main category name for the errors that are gathered is termed ‘error’; and all errors, having different names, are the subcategories of this category. Some of these names come from the domain knowledge; for example, ‘arithmetic’ and ‘sign’ errors. However, new names are required for new types of errors that are not in the domain language; for example, 'purposeful' error is one of these names developed for naming one of the subcategories of error.

The generated categories and/or subcategories are utilised to form abstraction. According to Elo and Kyngäs (2008) the process of abstraction takes the following steps: (1) generating categories that reflect the general description of the subject matter using the categories names that describe the characteristics of the subject the categories refer to, (2) iterating the abstraction process till the representation of the general description of the phenomena is satisfactory and (3) grouping the subcategories that share similar specifications under higher order categories. These processes have been applied to form abstraction using categories and subcategories of solution strategies and errors; sections 5.4.1 and 5.4.2 will present these categories and subcategories in addition to their descriptions in both figure and table formats.

5.4 Qualitative scripts content analysis result presentation

The reporting tools to present the analysis results are (1) figures displaying the categories and subcategories relationships (2) tables of definitions (3) raw data from student scripts, and (4) analyses of student workings. The figures that present the categories and subcategories show the names of the entities and the hierarchical relationships between these entities. Their explanations are given in the tables. Each entity's functionality or role as well as relationships between entities are also detailed in the tables. The tables, furthermore, contain examples of creating constructs and models using categories and subcategories. The constructs and models are used to explain the characteristics of student workings in terms of providing knowledge that assist in understanding the current context. The visual display of the original student scripts alongside the analysis' documents show the raw data and the analysis process. That display, along with the results and examples of the creation of categories and/or subcategories would add strength to the validity, credibility and reliability of the analysis process as well as its outcome.

In section 5.4.1 all the categories and subcategories and their meaning are presented. In section 5.4.2 examples of selected student scripts and the analysis' documents will be shown and discussed; this section also includes the previous research outcomes that are comparable with some of the findings in the analysis of the scripts.

5.4.1 Categories and subcategories presentation.

The two types of categories and their subcategories, and their meanings will be presented in this section. The top down approach that is shown in Figure 5.18 and Figure 5.19 follows Rose's, (2015) categories and subcategories presentation. The subcategories and their meanings will be used in answering the qualitative research questions. When answering the two qualitative research questions the details of the sections that contain their answers will be stated. However, sections such as 5.1, 5.2, and 5.3 also contribute information or knowledge in answering these questions. In addition, the data that has been produced by qualitative and quantitative analyses of the scripts contain data and/or knowledge that is more than that is required to answer the research questions. These research questions are general, when compared with the amount of information that is produced by the qualitative and quantitative analyses of scripts. The main purpose of the analyses of the scripts is not to only answer these questions but also to produce knowledge that will be utilised in producing the design science research artefacts.

The first qualitative question is:

What types of solution strategies have been used to solve problems?

The solution strategies category and its subcategories contain the answer to this question. Figure 5.18 shows these strategies, which are explained, in Table 5.2. The details of the Table 5.2's columns contents meanings are listed in Table 5.1. There are 12 types of subcategories of solution strategies displayed in Table 5.2. Further explanations of these subcategories and their meanings are given in appendix B.

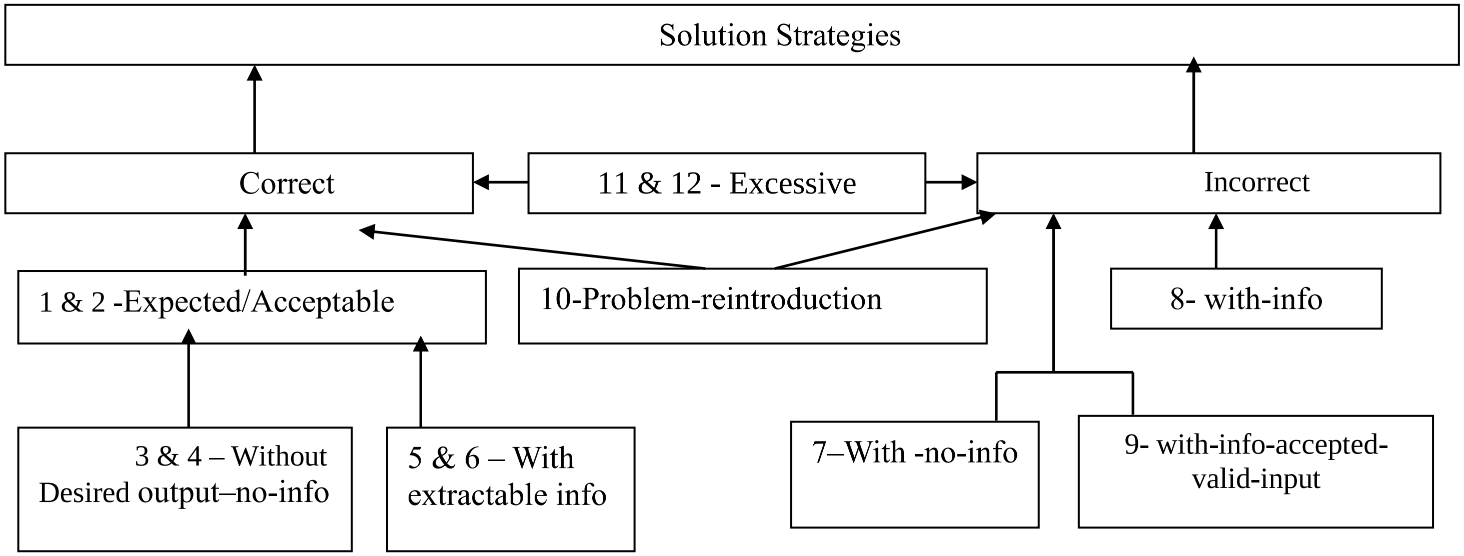


Figure 5.18 Solution strategies subcategories

Table 5.1 Meanings of the columns' contents in Table 5.2

Column names	**Column content meanings**
Subcategories' numbers	The number that is used to refer to the subcategories in Table 5.2
Names/types of solutions strategies and subcategories	The name or type of the subcategory, which represents the solution strategy and it is the main subject in its row. The rest of the columns (entities) in the row describe the purposes of this subcategory to perform tasks in student workings and the consequences of performing the tasks.
Desired outputs	The outputs of applying the subcategory (solution strategy) in the subcategory column. A tick means that the output of the application of subcategory solution strategy is the desired output. An X means an error or multiple errors is/are present when the subcategory of solution strategy is used incorrectly.
Errors	'Errors' is the main category name representing all the errors subcategories that might occur during the application of the subcategory that is displayed in the subcategories column. An X indicates that no error is present in the process of applying the solution strategy, and a tick shows the presence of one or multiple types of errors.
Valid inputs	The desired output or step's output that is produced as the result of the application of the current solution strategy subcategory is accepted as the valid input in the next step's process. A tick means that the desired output can be used as the valid input in subsequent steps. An X indicates that the current step's desired output is not accepted as the valid input in the subsequent steps. The presence of tick and X at the same time shows the possibility of the acceptance or the rejection of the desired output as the valid input in the next steps.
Awarded	This column represents the consequences of the selected or presented solution strategy's subcategory application or its presence. A tick indicates that a mark has been awarded to the student working. However, the X value reflects the provided solution strategy subcategory is not awarded any mark; the presence of tick and X at the same time indicates the possibility of awarding of the partial marks.

Note. This table contains the meaning of the columns in Table 5.2 that contains the subcategories that have been created from the result of the analysis of student working. *Note the value of N/A in all columns of* Table 5.2 *has the same meaning. This is to say that the value of N/A for all column shows the irrelevance of this particular attribute (column) for the solution strategy's subcategory that is under discussion, in the row.*

Table 5.2 Solution strategies subcategories and their meanings

Subcategory types numbers	**Solutions strategies' subcategories names (types)**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
1	Correct expected	✓	✕	✓	✓
2	Correct acceptable	✓	✕	✓	✓
3	Correct expected without desired output	**N**/A	**N**/A	**N**/A	✓X
4	Correct acceptable without desired output	**N**/A	**N**/A	N/A	✓X
5	Correct expected with extractable info	✕	✓	✓✕	✓✕
6	Correct acceptable with extractable info	✕	✓	✓✕	✓✕
7	Incorrect unacceptable/unrelated with no info	**N**/A	**N**/A	**N**/A	**N**/A
8	Incorrect unacceptable/unrelated with info	✕	✓	✕	✕
9	Incorrect unacceptable/unrelated with info and accepting the solution produced	✕	✓	✓	✕
10	Correct solution strategies that reintroduces previously solved problems	✓	✓	✓	✓X
11	Excessive/Unnecessary Work with error and/or info	✕	✓	✕	✕
12	Excessive/Unnecessary Work with no error or info	✕	✕	✕	✕

The second qualitative research question is shown next.

What types of errors have students made during solving problems?

Figure 5.19 shows the subcategories of errors representing the answer to this question. The meanings of 11 types of errors are displayed in Table 5.3; further explanations of these subcategories are also given in appendix B.

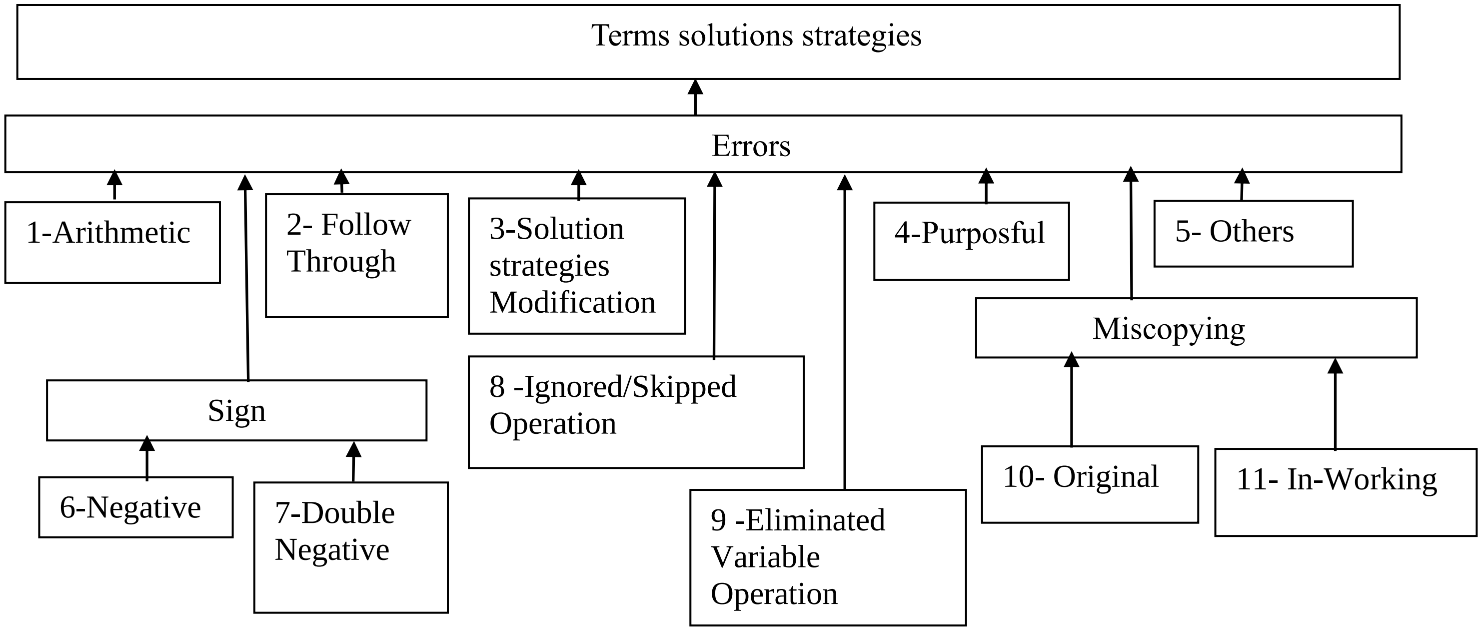


Figure 5.19 Errors subcategories

Table 5.3 Errors subcategories' entities and their meanings

Categories Number	Category and Subcategories names (types)	Meanings
	Error	'Error' is the category that indicates all the subcategories of error.
1	Arithmetic error	'Arithmetic error' is the error arising from the incorrect mathematical computation excluding negative sign, double negative sign and purposeful errors.
2	Follow through error	'Follow through error' is an error or multiple of errors that have occurred in previous steps and transmitted to subsequent steps. The presence of follow through error in the subsequence steps changes steps results but should have no effect in awarding marks for workings.
3	Solution Strategies Modification	The 'solution strategies modification error' is the error in which the correct solution strategy is changed into the incorrect solution strategy. This may produce the incorrect answer and may or may not affect the awarded mark.
4	Purposeful error	The 'purposeful error' is the error that students intentionally make to produce the intended correct result. Here students use the incorrect solution strategy that produces a partially correct answer but introduce a partly incorrect solution. However, students ignore the incorrect part of the answer and use the correct part as the correct solution that represents the complete outcome of that step. Further explanation of this subcategory is given in section 5.4.2.2.
5	Other error	This is an unexplainable type of error within a step affecting the operation in the step.
6	Negative Sign Error	This is the error that occurs when the problem solvers ignore or forget numbers being negative when applying a correct solution strategy to perform a task but produces the incorrect output to that step.
7	Double Negative Sign Error	The double negative sign error is the error that is not a negative sign error, but it is caused by the presence of the subtraction operation using the negative numbers in which the problem solvers fail to convert the double minus sign to the positive sign. Further details of this subcategory are given in section 5.4.2.3.

Categories Number	Category and Subcategories names (types)	Meanings
8	Ignored/Skipped Operation error	The 'ignored operation' error is the error that occurs when students ignore the appropriate mathematical operation on the part of a task or process. For example, eliminating a term from an equation, leading to an incorrect answer. It resembles the purposeful error, but it is different. Further detail of this subcategory is given in section 5.4.2.1.
9	Eliminated Variable Only Operation	This error occurs when the correct solution strategy is used to eliminate a term from an equation; however, the column in the equation that contains the term to be eliminated is the only part in the equation that is affected by the solution strategy application. The other columns, in the equation, remain unchanged. This is selectively applied to yield only the desired result. For example, in an element elimination process, a student performs (3, 2, 4, 5) – (3, 1, 3, 4) producing 0, 2, 4, 5 instead of 0, 1, 1, 1. The analysis of the student working displays performing the subtraction operation on only the element on the left hand column leaving other elements an affected by the operation.
10	Original Question miscopying	This may lead to a correct strategy being applied to a different problem to the one set in the examination.
11	In-working miscopying	It is the error that occurs due to miscopying any section of the working: numbers, symbols, coefficients, and mathematical signs of these entities during the solution. This error has similar effects as the original question miscopying error, although the change in the question's solution only applies to a particular part of the solution.

The next section demonstrates the analysis process using samples from student workings and the analysis' documents; the section also reviews previous research findings to show that the student workings analysis results are valid, credible and reliable.

5.4.2 Roles of Student scripts and analysis documents in categories and subcategories presentation.

The attributes that ensure the acceptability of the created categories and subcategories have been explained in sections 5.1 and 3.3.3. As explained in section 5.1, one of the attributes are to demonstrate the acceptability of the output of research investigation through establishing the relationships between the input, which is the raw data that has been investigated, and the output, which is the investigation result. In section 5.2, some of the relationships between student scripts, the categorisation of the findings in the analysis, and the complexity of understanding some of the data in the scripts have been identified and discussed using the scripts and the analysis' documents as examples. Figure 5.20 shows the relationships between the scripts, the analysis process, and the analysis' documents. There is a two-way relationship and the figure shows the iterative nature of the analysis process.

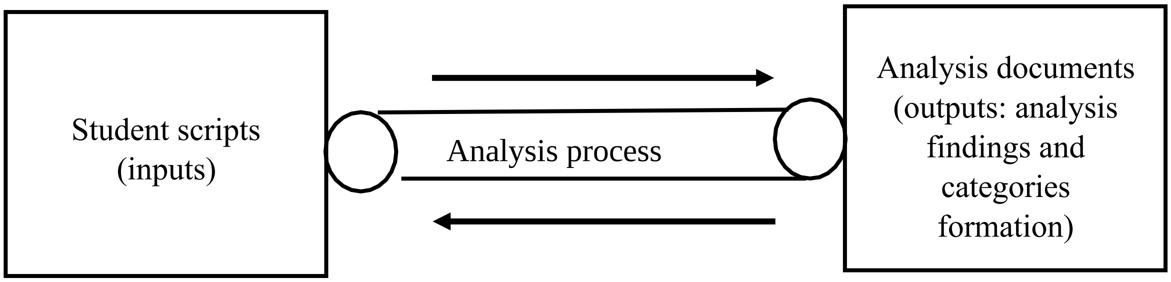


Figure 5.20 Inputs, pipeline and outputs relationships

The selected scripts and the documents' contents will be used to demonstrate how categories and subcategories have been formed from the findings in the analysis. The entire categories and subcategories and their meanings presentation is in section 5.4.1. The main task in this section is to present examples, as samples, using student scripts and the analysis' documents to show the process of producing categories. The purpose of performing this task is to demonstrate the validity of the outputs of the analysis by showing the analysis process that is linked to the raw data which is the student scripts.

5.4.2.1 Ignored operation error subcategory formation demonstration.

In this section, the subject to be addressed is the *ignored operation error* subcategory. The student working section, concerning this subcategory, is shown in Figure 5.22 while the complete workings of the student is displayed in Figure 5.21. The detail of the analysis is displayed in Figure 5.23. The relevant section of the document is bounded by the two horizontal lines. The analysis of the student workings shows that the intention was to eliminate the x term from equation two using R2 – 2R1 solution strategy. The correct application of the strategy produces 0, 1, 3, 2. However, the student's answer is 0, 5, -3, 10. This indicates that the subtraction was only applied to the column on the left, which has the value of zero. Values in other three columns have remained unchanged. These three columns have the same values as the three corresponding columns that are shown in R2. Thus, the creation of the ignored operation error category is based on applying the solution strategy only to one element of the matrix while ignoring the other elements; and the name of the subcategory is based on this understanding, which is ignoring the application of the solution strategy to some of the elements.

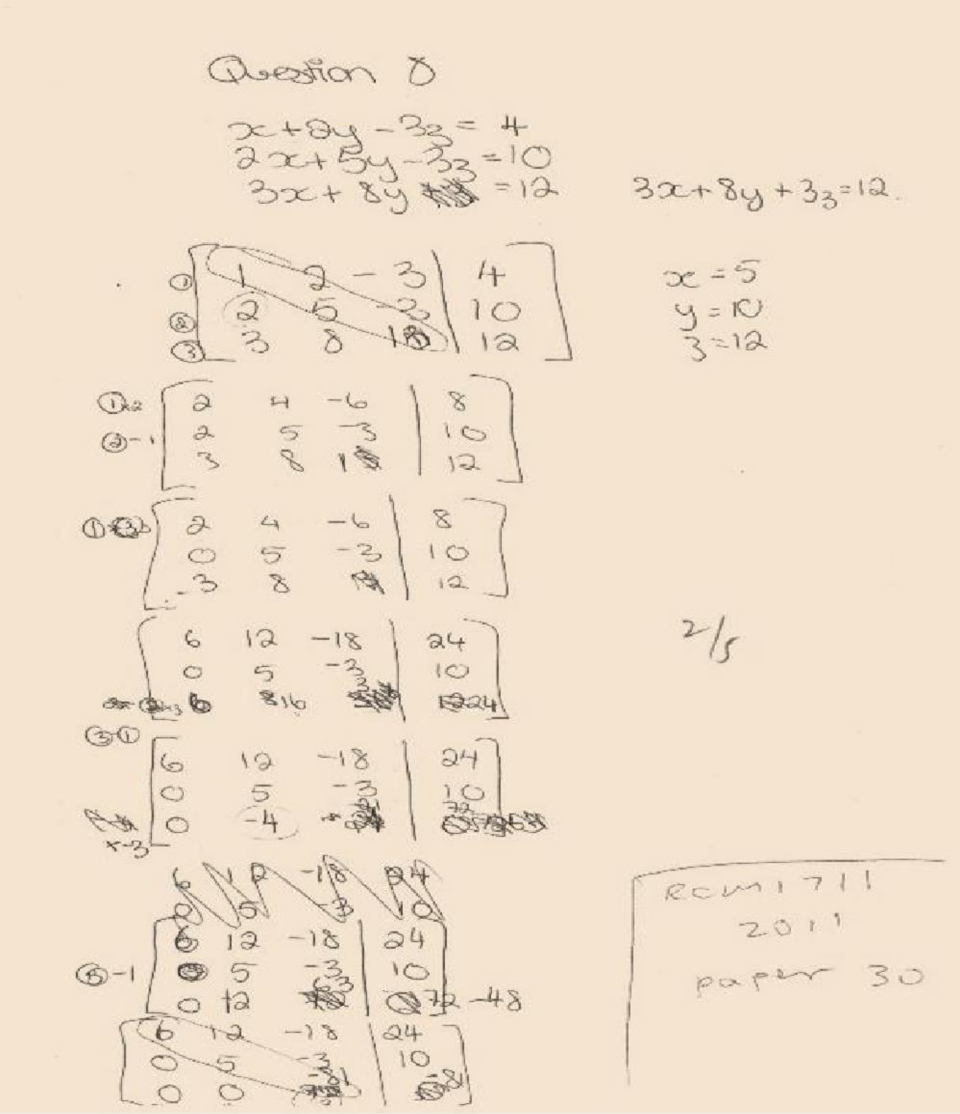


Figure 5.21 Ignored operation error subcategory in student working

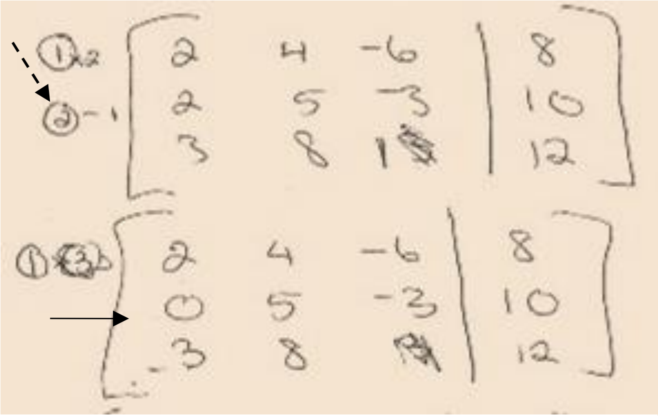
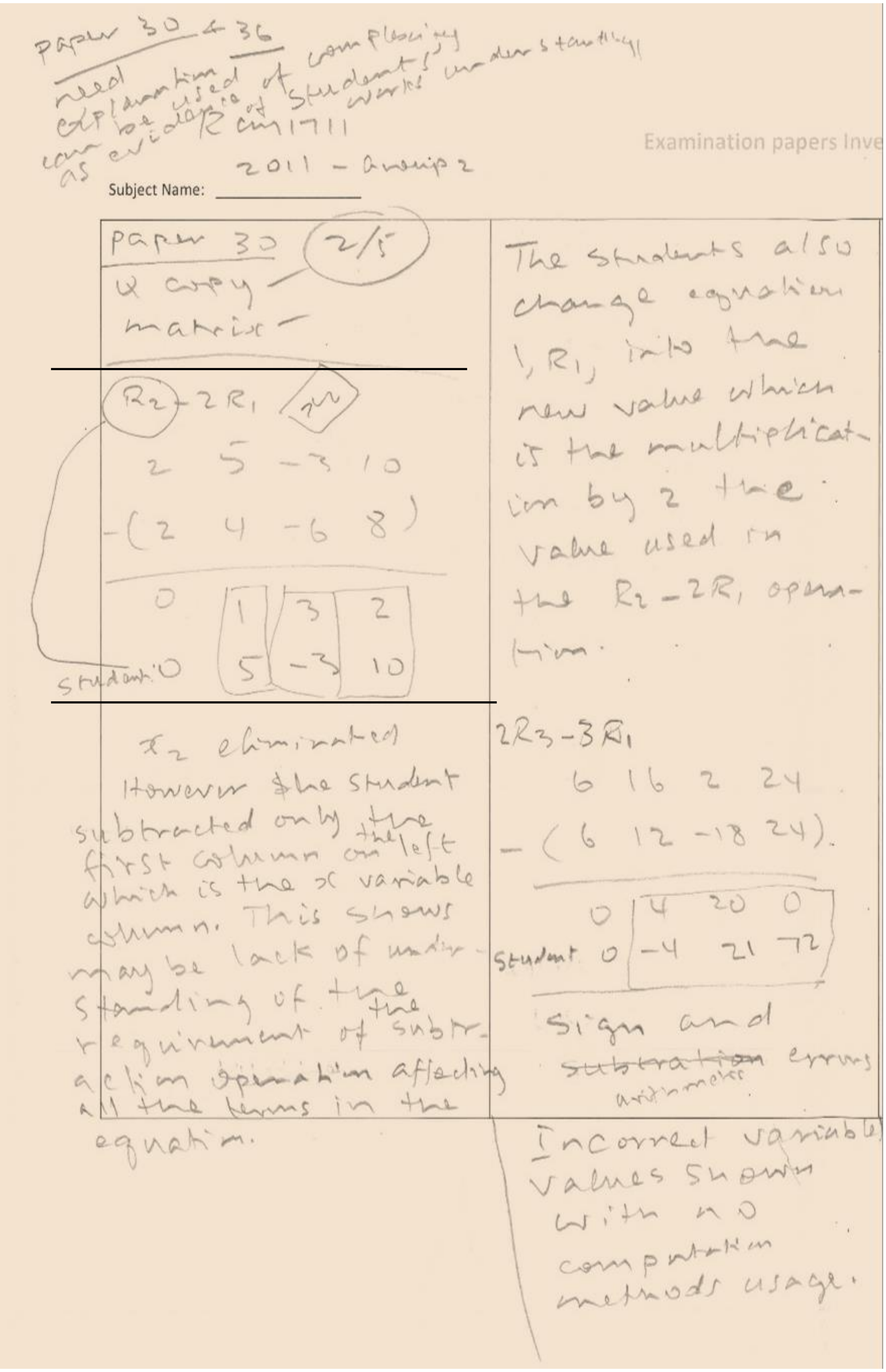


Figure 5.22 Section of Figure 5.21



paper 30 & 36

need explanation of complexing students' understanding

can be used as evidence of students' works

R cm1711

Examination papers Inve

2011 – Group 2

Subject Name: ____________

paper 30 (2/5)

Q copy matrix –

$R_2 - 2R_1$

2 5 −3 10

−(2 4 −6 8)

0 1 3 2

student 0 5 −3 10

The students also change equation 1, R_1, into the new value which is the multiplication by 2 the value used in the $R_2 - 2R_1$ operation.

x_2 eliminated

However the student subtracted only the first column on the left which is the x variable column. This shows may be lack of understanding of the requirement of the subtraction operation affecting all the terms in the equation.

$2R_3 - 3R_1$

6 16 2 24

−(6 12 −18 24).

0 4 20 0

student 0 −4 21 72

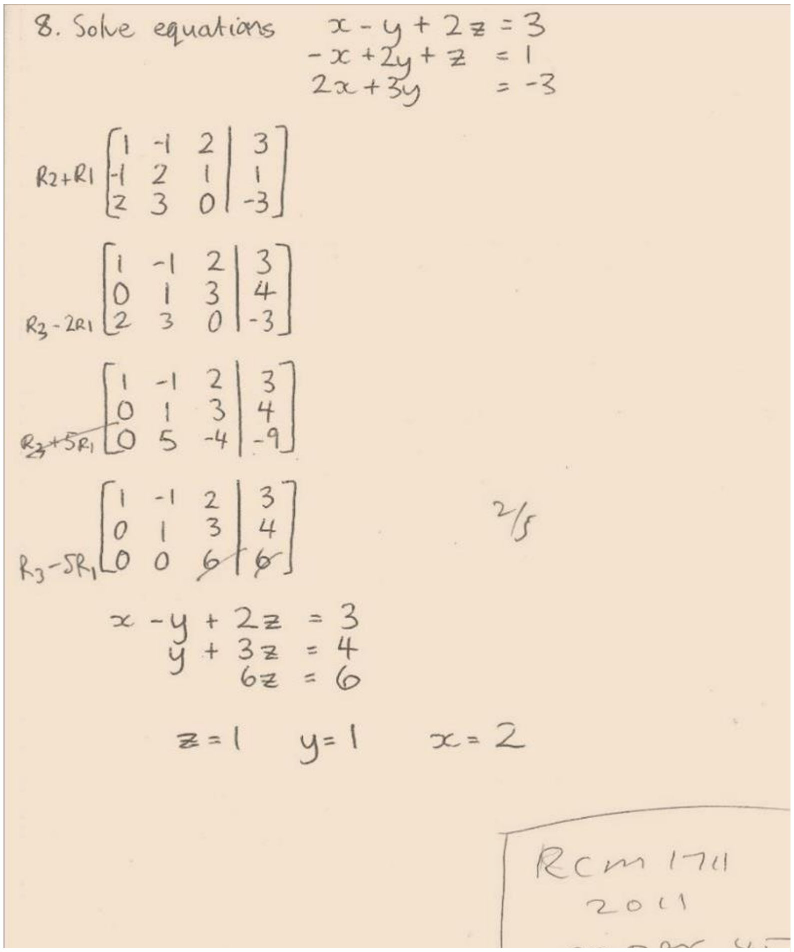
Sign and ~~subtraction~~ arithmetic errors

Incorrect variable values shown with no computation methods usage.

Figure 5.23 Analysis of student working forming ignored operation error subcategory

5.4.2.2 Demonstration of multiple subcategories formation.

In this section, multiple examples of subcategories formation by demonstrating the analysis of student workings will be discussed. In student workings, which are presented in Figure 5.24, there are two correct expected solution strategies that have been applied to achieve the intention of the student. This intention is to eliminate the term x from equations two and three using R2 + R1 and R3 – 2R1 strategies. The analysis' document, displayed in Figure 5.25, shows that the word 'expected' has been added as a comment next to the x in the second and third equations. This word means that their elimination has been performed as expected. The reason for giving the word expected as category's/subcategory's name for this



8. Solve equations $x - y + 2z = 3$

$-x + 2y + z = 1$

$2x + 3y = -3$

$$\begin{array}{l} \\ R_2+R_1 \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ -1 & 2 & 1 & 1 \\ 2 & 3 & 0 & -3 \end{array}\right]$$

$$\begin{array}{l} \\ \\ R_3-2R_1 \end{array}\left[\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 0 & 1 & 3 & 4 \\ 2 & 3 & 0 & -3 \end{array}\right]$$

$$\begin{array}{l} \\ \\ R_3+5R_1 \end{array}\left[\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 0 & 1 & 3 & 4 \\ 0 & 5 & -4 & -9 \end{array}\right]$$

$$\begin{array}{l} \\ \\ R_3-5R_1 \end{array}\left[\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 6 & 6 \end{array}\right]$$

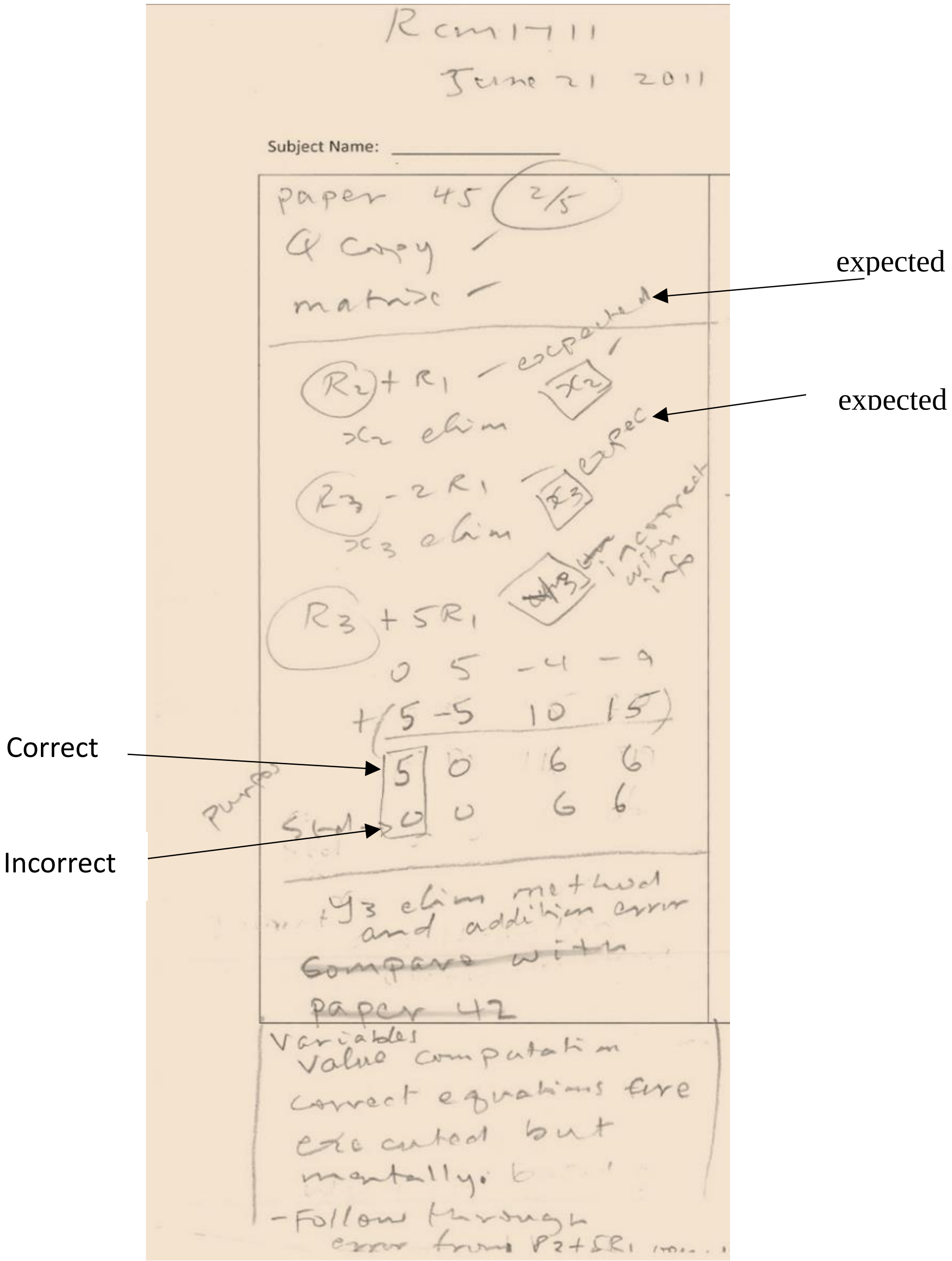
2/5

$$\begin{aligned} x - y + 2z &= 3 \\ y + 3z &= 4 \\ 6z &= 6 \end{aligned}$$

$z = 1$ $y = 1$ $x = 2$

Rcm 1711
2011

Figure 5.24 Multiple subcategories in student working strategy has been given in section 5.3. The other two subcategories that were formed from the findings that are seen in the document are ‘incorrect with info’ solution strategy and the ‘purposeful’ error.



Rcm 1711
June 21 2011

Subject Name: ________________

paper 45 2/5
Q copy ✓
matrix ✓

R_2 + R_1 ✓ expected ✓ x_2
x_2 elim

R_3 - 2R_1 ✓ x_3 expec
x_3 elim

R_3 + 5R_1 ~~x_3~~ 3 incorrect with info

0 5 -4 -9
+(5 -5 10 15)
5 0 6 6
purpos
Stud → 0 0 6 6

y_3 elim method and addition error
~~Compare with paper 42~~

Variables value computation
correct equations are executed but mentally.
- Follow through error from R2+5R1 ...

Figure 5.25 Detection of multiple subcategories in student workings

The origin of these two subcategories are the R3 – 5R1 strategy that is applied to eliminate the term y from the third equation. The exploration of the student workings of this section presents the 0, 0, 6, 6 as the answer for the application of the strategy. However, the correct output for applying this strategy is 5, 0, 6, 6. In this example, the incorrect strategy has produced the incorrect answer. The student purposefully introduced an incorrect x term value to zero, ignoring the consequence, which is the reintroduction of the x term; because this term has already been eliminated in the previous task that was performed using R3-2R1. Consequently, the 'incorrect with info' solution strategy subcategory type is created and its meaning is given in Table 5.1 and Table 5.2. The other subcategory is the 'purposeful' error. The meaning of this subcategory is covered in Table 5.3. Furthermore, the two comments that are written next to the two mentioned subcategories, as shown in the document, demonstrate that the analysis process being extensive and deep enabling the production of data or knowledge that can be abstracted and claimed as theory. Hence, the visual display of the student workings alongside the analysis' documents enhances the validity of the output of the analysis.

5.4.2.3 Double minus sign error subcategory formation demonstration.

In the findings of the analysis of student scripts, different types of arithmetic errors have occurred. The focus in this discussion is on the operations that involve negative numbers. Bellamy (2015) states that negative numbers are applied in limited scopes such as scale and temperature. This limitation minimises the exposure to the operations that involve negative numbers (Prather & Alibali, 2008). The lack of "perceptual referents" (Blair, Rosenberg-Lee, Tsang, Schwartz, & Menon, 2012, p. 1) children have to put extra effort in grasping the negative number concept (Booth, Barbieri, Eyer, & Paré-Blagoev, 2014; Fuadiah & Suryadi, 2017); it is also difficult to associate negative numbers with a space (Kilhamn, 2011). These problems are amongst those that make the operations with negative numbers more challenging than operations that contain only positive numbers (Prather & Alibali, 2008).

The complexity of the operations add confusion when using minus signs both for negative numbers and for subtraction operations as it is used for both purposes. According to Booth et al. (2014), (Kullberg, 2010) and Vlassis (2004) students could not differentiate between the two purposes of the minus sign and believed that it is utilised in subtraction operations only. The presence of negative numbers in operations could add greater difficulty for some students than other factors in the equations structure (Kilhamn, 2011). One of the most challenging issues when a minus sign is used simultaneously for number negativity and for a subtraction operation. Here a double minus sign is applied (Gallardo, 1995), for example, (-(-b)). Double minus signs are very badly managed by students as Gallardo stated, “the percentage of correct answers in this case was 20%” (p. 4). Nicolas and Liora (1994) and Kilhamn (2011) identified the seriousness of “the detachment of a number from the preceding minus sign” (Nicolas & Liora, 1994, p. 76) as being high, and in the study conducted by Cangelosi, Madrid, Cooper, Olson, and Hartter (2013), the negative sign error in exponential expressions was persistent. The demonstration of two of the problems will be shown next using the analysis of one of student workings, the analysis process and the result of the analysis.

In the manifestation of the double minus sign error, in the analysis of the student workings and the result of the analysis, the selected student working sample is shown in Figure 5.26 and the analysis detail of this working is similarly shown in the analysis’ document in Figure 5.27. The entire student workings involve the subtraction operations containing negative numbers. In the document, the first task that has been analysed is the R2-2R1 solution strategy operation that is intended to eliminate the x term from the second equation. The double minus error occurs in the -3 - (-6) operation, which is presented in the second column from right. The student's result for the given operation is -3 while the correct output should be 3. The second problem is displayed in R3 - 3R1 strategy. The second column, from right, presents the 1 - (-9) operation. The correct answer for this operation should be 10; however, the student has given the numerical value 8 as the answer. Thus, the above two examples have displayed the problem of double minus sign being persistent because the student has repeated the same error in the same setting. The student error persists to the next operation which is applying 2R2 – R3 strategy to eliminate the term y from the third equation. The result of the operation shows the existence of two errors; however, from the observation it seems the answer the student has incorrectly computed - 8 - 6 to be 2, shows the detachment of the number 8 from the preceding minus sign. Another error, which is 4 – 0 equalling 0 has no obvious explanation.

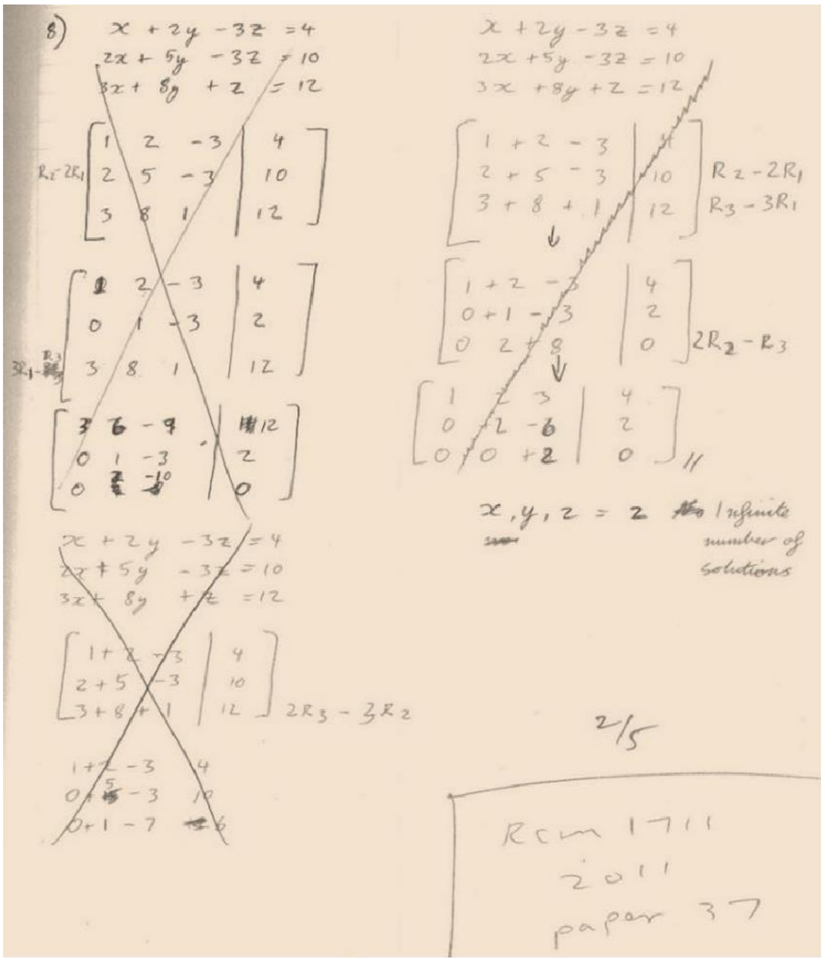


Figure 5.26 Double minus sign error in student workings

In summary, there are many errors that can be associated with the negative number's usage; some could be explained, and some could not. The important issue is the acknowledgement of the existence of these multiple errors and accepting them as the factual concepts that advance the knowledge in the domain. Thus, from the results of the analysis and the previous publications in the subject, the double sign minus error occurrence is not new confirming or validating the findings of this subcategory in this research.

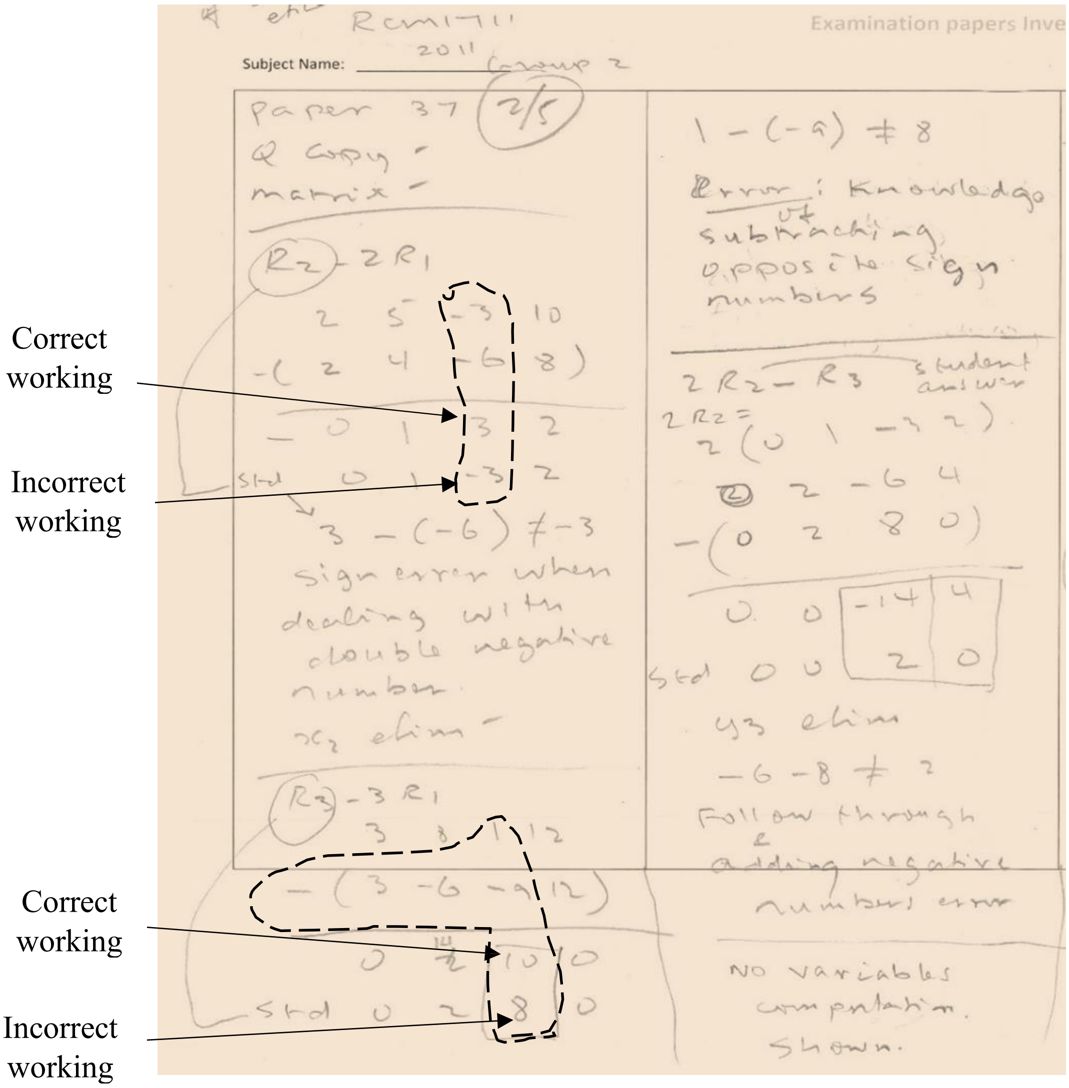


Figure 5.27 Analysis of double minus sign error in student workings

The qualitative data has been produced through iterative investigation of the scripts that has involved the interpretation and discussion of the qualitative data. The interpretation and discussion of the data have been performed during the analysis of the scripts, when examining findings to form categories and subcategories, and when forming the questionnaire as well as during the quantitative analysis of the scripts. Furthermore, sections 5.1, 5.2, 5.3, 5.4.1 and 5.4.2 have also contributed to the discussion and interpretation of the data in the qualitative analysis of scripts. Therefore, it is not necessary to discuss and/or interpret the output of qualitative analysis of the scripts again in a separate section.

5.5 Quantitative Student workings Survey Analysis Results

The quantitative research designs and the details of the selected approach and methods have been discussed in section 4.5. Figure 4.4 has shown steps in the quantitative data analysis of student scripts. The first, second and third steps have been explained and implemented in section 4.5.5. The fourth step in the figure is answering research questions. In this research, mixed methods methodology has been used when analysing the student workings. The qualitative research questions have been answered during the presentation of categories and subcategories of solution strategies and of errors. In this research, the roles of the quantitative analysis of scripts are to answer questions that add more explanations to the qualitative data and to support answers that are given for the qualitative questions.

The fourth step and the fifth step which is ‘presenting findings’ (both steps are quantitative) will be reported and discussed in the subsequent subsections.

The following format will be used to report and discuss the result of the analysis.

1. Explain the purpose of collecting the data.
2. Display a question or questions, which collect(s) the quantified data for analysis.
3. Display the question that determines what to analyse of the collected data (this is the same as the quantitative research questions).
4. Display the analysis result in tabular and graphical formats.

5. Explain and/or discuss the presented data.

In the data that has been collected for the quantitative analysis, the 'missing' data value was shown as 'system'. In the analysis of scripts, the missing data is interpreted in different ways. If the analysed data is about errors in the student working, the presence of 'system' as the value of the 'missing' data indicates that the student working does not contain that type of error. The absence of the error data is either due to the student working being awarded full marks or the student did not attempt to solve the problem. However, if the missing data is about a correct solution strategy to be employed in performing a task, the presence of the missing data indicates that the student working did not contain that particular solution strategy, which shows the lack of this knowledge by the student.

The data in the presentation contains answers to the quantitative research questions, which will also be referred to, in some sections, as the analysis questions. Three main quantitative research questions will be answered in the following sections. Two questions will be answered using data in solution strategies while the other question will be answered using the data in the student errors.

5.5.1 Solution strategies.

There are two main tasks in the scripts that have employed solution strategies. These are elimination of matrix elements and computation of the values of variables (unknowns).

In this section, the data that will be presented and discussed will display information about the order of operations and types of solution strategies that have been used in these two tasks.

5.5.1.1 Order of operations quantitative analysis report and discussion.

The purpose of collecting the order of operations data is to find out whether students have used the solution strategies in the expected order when solving the given equations. The formation of an upper triangular matrix in a Gaussian elimination, from a general 3x3 matrix, requires the elimination of three elements. Students are expected to eliminate elements a_{21}, a_{31} and a_{32} as shown in Figure 5.28. However, to avoid subscripts, matrix elements will be indexed by rows, as shown in Figure 5.29; and these numbers will be used to indicate positions in the matrix.

$$\begin{bmatrix} a11, a12, a13 \\ a21, a22, a23 \\ a31, a32, a33 \end{bmatrix}$$

Figure 5.28 Expected elements to be eliminated

1, 2, 3
4, 5, 6
_7, 8, 9

Figure 5.29 Index representation of elements positions

The question to collect the data to be analysed was:

Which element has been eliminated first from a general 3x3 matrix?

The same question was modified with the alteration of the 'first' word to 'second' and 'third' respectively when collecting data to discover the second and third eliminated elements.

The analysis question is shown next.

Question one:

How many students have eliminated elements 4, 7, and 8 in Figure 5.29 *sequentially to create an upper triangular matrix in a Gaussian elimination?*

The results of this analysis question are shown in Table 5.4, Table 5.5 and Table 5.6. In these tables the category value, which represents the eliminated element, varies between 1 and 9 inclusive; these numbers represent any element in 3x3 matrix. However, the three tables contained some or all the elements 4, 6, 7 and 8. As explained above these numbers refer to the element's positions in the matrix. These are the elements that students have selected when they have eliminated and/or have attempted to eliminate elements from the matrix.

The output of the analysis in Table 5.4 has resulted from the analysis of the data that was collected to find out the number of students who have selected the fourth element as the first element during element elimination tasks. This table shows the selection of the fourth, sixth and seventh elements and the number of students who have selected each element.

Table 5.4 Sequence of operation (first element's elimination)

		Frequency	Percent	Valid Percent
Valid	4	151	94.4	98.1
	6	1	.6	.6
	7	2	1.3	1.3
	Total	154	96.3	100.0
Missing	System	6	3.8	
Total		160	100.0	

able 5.5 shows the result of the analysis for the data that was collected to uncover the number of students who have shown the intention of eliminating the seventh element as the second element in the process of making the triangle. The data in this table shows the number of students who have selected the seventh and the eighth elements as well as the absence of some student workings data, which is reported as missing.

Table 5.5 Sequence of operation (second element' elimination)

		Frequency	Percent	Valid Percent
Valid	7	147	91.9	98.7
	8	2	1.3	1.3
	Total	149	93.1	100.0
Missing	System	11	6.9	
Total		160	100.0	

The result of the analysis that is shown in Table 5.6 is for the data that is collected to find out the number of students who have selected the eighth element as the third element to eliminate. In the table the students have selected the fourth, seventh and the eighth elements and some students did not show any selection which is shown in the table as the 'missing' data.

Table 5.6 Sequence of operation (third element' elimination)

		Frequency	Percent	Valid Percent
Valid	4	2	1.3	1.5
	7	1	.6	.8
	8	130	81.3	97.7
	Total	133	83.1	100.0
Missing	System	27	16.9	
Total		160	100.0	

The exploration of data, for example, in Table 5.4 shows that 151students have selected the fourth element as the first element to be eliminated. The selection of elements for elimination does not mean that all students have eliminated the elements from the matrix. What it conveys is that students have used a solution strategy or strategies that might or might not correctly eliminate the element.

Table 5.4, Table 5.5 andTable 5.6 also show missing values. If these missing values are excluded from the analysis, many students have chosen the correct expected elements for elimination. The data shows that the fourth element was selected, as the first element, by

98.05%, (151/154 ×100), of students, the seventh element was selected, as the second element, by 98.66%, (147/149 × 100), and the eighth element was selected, as the third element, by 97.74%., (130/133 × 100). This shows that large number of students have shown the understanding of eliminating elements in the expected (correct) order.

The next section will present data that shows types of solution strategies that have been employed in eliminating elements as well as their frequency.

5.5.1.2 Element elimination solutions strategies quantitative analysis report and discussion.

Table 5.2 and Figure 5.18 show 12 types of solution subcategories that have been formed from the findings of the qualitative analysis of student scripts. The types of solution strategies that students have used in their workings are the principal concerns of the assessors. In other words, they are used to indicate levels of students' knowledge of a domain. The purpose of analysing types of solution strategies is to uncover different types of solutions strategies that have been used when eliminating elements. To answer this question the following question was used to collect data:

Which elimination methods have been used to eliminate elements to form an upper triangular matrix in a 3x3 matrix?

The following question has been used to analyse the collected data.

Question two:

*What types of solution strategies have been used to eliminate the fourth, the seventh and the eighth elements from the matrix? (*See Figure 5.29 for these elements)

Table 5.7 contains the types of subcategories of solution strategies and the numbers (1 to 12 inclusive) that refer to them. The result of the analysis that shows the subcategories of solution strategy types (represented by numbers) and the number of students who have used these strategies is displayed in Table 5.8, Table 5.9 and Table 5.10. The details of the meanings of these subcategories, are given in Table 5.1 and appendix B.

Table 5.7 Solution strategies types and their reference numbers

Subcategory types numbers	Solutions strategies' subcategories
1	Correct expected
2	Correct acceptable
3	Correct expected without desired output
4	Correct acceptable without desired output
5	Correct expected with extractable info
6	Correct acceptable with extractable info
7	Incorrect unacceptable/unrelated with no info
8	Incorrect unacceptable/unrelated with info
9	Incorrect unacceptable/unrelated with info and accepting the solution produced
10	Correct solution strategies that reintroduces previously solved problems
12	Excessive/Unnecessary Work with no error or info

Table 5.8 shows different types of solution strategies which include types 1, 2, 3, 8 and 11, which have been used to eliminate the fourth element in the matrix. The number of students who have used each type of the strategy is also shown.

Table 5.8 Type of solution strategies for eliminating fourth element

		Frequency	Percent	Valid Percent
Valid	1	154	96.3	96.3
	2	3	1.9	1.9
	3	1	.6	.6
	8	1	.6	.6
	11	1	.6	.6
	Total	160	100.0	100.0

Regarding the solution strategy to eliminate the seventh element, Table 5.9 shows that students have used the solution strategies types 1,2,3,7 and 8. The table also includes the missing data that indicates the absence of some of student workings regarding the elimination of the seventh element.

Table 5.9 Type of solution strategies for eliminating second seventh element

		Frequency	Percent	Valid Percent
Valid	1	147	91.9	93.6
	2	4	2.5	2.5
	3	2	1.3	1.3
	7	2	1.3	1.3
	8	2	1.3	1.3
	Total	157	98.1	100.0
Missing	System	3	1.9	
Total		160	100.0	

To determine the solution strategies types that have been used to eliminate the eighth elements the result of the analysis is shown in Table 5.10. The result shows that the strategies types that have been used are 1, 2, 3, 8 and 9. The other information in this table is the missing data.

Table 5.10 Type of solution strategies for eliminating eighth element

		Frequency	Percent	Valid Percent
Valid	1	115	71.9	83.3
	2	1	.6	.7
	3	6	3.8	4.3
	8	11	6.9	8.0
	9	2	1.3	1.4
	10	3	1.9	2.2
	Total	138	86.3	100.0
Missing	System	22	13.8	
Total		160	100.0	

The exploration of Table 5.8, Table 5.9 and Table 5.10 show that values (numbers) that have been referred to by the ‘valid’ column represent types of subcategories of solution strategies. In the three tables the most prevalent type of strategy is type 1. The 1 refers to the 'correct expected' type of solution strategy. The number of students who have selected this strategy to eliminate elements numbers 4, 7 and 8 are shown as 96.3%, 91.9% and 71.9% respectively.

The results of the comparison of the analysis of order of elimination that has been shown in Table 5.4Table 5.5 and Table 5.6 and the analysis of types of solution strategies that has been displayed in Table 5.8, Table 5.9 and Table 5.10 show some differences in the number of students who have selected elements to be eliminated and those who have chosen particular types of solution strategies to eliminate these elements. For example, the data that is in Table 5.4 shows the number of students who have selected elements for elimination as well as six students' data reported as missing. By contrast, Table 5.8, which shows types of solution strategies to eliminate the fourth element, does not include missing data to be reported.

This difference is attributed to (1) some types of solution strategies performing no task; and any data in the student workings that contains this type of solution strategies are only included in the analysis to discover types of solution strategies and not in the analysis that informs the order of operations. For example, if this type of strategy is provided by students during the process of eliminating elements, the data that shows the selection of this strategy by students will not be included in the analysis to find out order of eliminating elements. (2) There were also five students who have used total elimination method (eliminating all matrix elements of the main diagonal) to compute for the values of variables. These students' data does not conform to the order of operations that uses Gaussian method to form a triangle in a matrix. However, these student workings were analysed for two purposes: (1) to uncover types of solution strategies and (2) to gather errors from student workings. Hence, including these students' data in the analysis of order of eliminating elements is irrelevant. However, including these student' data in the analysis to uncover the types of solution strategies that have been used in eliminating elements is relevant. The types of solution strategies that have been used during the computation of values of variables will be presented and discussed in the next section.

5.5.1.3 Variables values computation solutions strategies quantitative analysis report.

There were three variables that students were required to compute their values. The same numbers to refer to types of solution strategies that have been used in eliminating elements have been used in referring to types of solution strategies that have been used to compute for the values of variables.

Table 5.11 displays the types of solution strategies that are used to compute the value of the z variable and the numbers of missing data. It also shows the number of students that have used each type of the strategy.

Table 5.11 z variable value computation strategy

		Frequency	Percent	Valid Percent
Valid	1	112	70.0	98.2
	3	1	.6	.9
	8	1	.6	.9
	Total	114	71.3	100.0
Missing	System	46	28.8	
Total		160	100.0	

The y variable is the second value to be computed. Table 5.12 shows the types of strategies, which are 1, 3 and 8 that have been used to find out the value of the variable. The table also displays the number of students who have participated and successfully computed the value of the variable as well as those who have attempted computing the value of the variable. In addition, the table includes the missing data.

Table 5.12 Variable y value computation strategy

		Frequency	Percent	Valid Percent
Valid	1	97	60.6	96.0
	3	2	1.3	2.0
	8	2	1.3	2.0
	Total	101	63.1	100.0
Missing	System	59	36.9	
Total		160	100.0	

The x variable was the last value to be computed. In Table 5.13, the result of the analysis includes types of solution strategies and the number of students who have used these strategies to solve and/or attempt to solve the value of the x variable. Furthermore, the result shows the missing value representing the number of students whose works are absent from the analysis.

Table 5.13 Variable x value computation strategy

		Frequency	Percent	Valid Percent
Valid	1	94	58.8	96.9
	3	2	1.3	2.1
	8	1	.6	1.0
	Total	97	60.6	100.0
Missing	System	63	39.4	
Total		160	100.0	

The examination of types of solution strategies that students have used to compute values for the variables shows a similar result to the type of solution strategies that have been used in eliminating elements. Most students who have attempted to compute values for the variables have used the ‘correct expected’ strategy type showing the knowledge of using correct solution strategies. In Table 5.11, Table 5.12 and Table 5.13 the missing data increases as students advance towards the end of solving the equations. For example, when computing the value of x variable, the missing data was 39.4% compared to 28.8% when computing the value of z variable. This shows that approximately 40% of students were not able to complete their works. This might indicate the aggregation of errors in the student workings preventing students successfully solving the equations. In the student workings, the last task to be performed is computing the value of x and this computation is dependent on the availability of y and z variables' values.

5.5.2 Errors types quantitative analysis report.

The basis for the quantitative errors data is the subcategories errors in the qualitative data. In this section answers to the quantitative research questions (analysis questions) with respect to the student errors are given in the analysis of these errors. In this section, the quantitative analysis of the following eight types of error are presented and discussed.

1. Arithmetic error.
2. Minus sign error.
3. Double minus error.
4. Purposeful error.
5. Ignored operation error.
6. Others error.
7. Equation arrangement error (Solution Strategies Modification).
8. Questions and in-working miscopy errors.

The purposes of collecting quantitative data about the student errors are to determine the frequency of each error's occurrence and to learn the places in the workings that produced these errors. When collecting the data, the same questions in the questionnaire have been repeatedly used when the same information is collected from different places of student workings. For example, during the elimination of the element 4 all types of errors or solution strategy types that had been used were recorded. The same type of information was collected when eliminating other elements and computing for values of variables. Further details of all the questions in the questionnaire are given in appendix A. However, as an example some of the questions that were used in collecting the arithmetic error are given in section 5.5.2.1.

When analysing the collected data, the same question or questions can be employed to analyse error data; because the aims of the analysis' questions are to produce the frequency of occurrences of these errors and to identify the places of the scripts that have produced them. Therefore, posing the following questions to all type of errors variables will suffice.

Question three:

What is the frequency of this type of error across all scripts and where, in the student workings, does this error appear?

This question will be used in analysing the data about the eight types of error that have been listed above. The result of the analysis for each error will be presented in separate tables. Only the data corresponding to arithmetic errors will be explained. The same explanation will apply to the contents of each table that displays the analysis of all types of errors.

Each of these tables consists of three columns. The leftmost column contains the places in the student workings where this error occurs. The second column contains the number of occurrences of this error in each section. The third column shows the relative frequencies of each place.

5.5.2.1 Arithmetic error.

The question that collected the occurrences of an arithmetic error when computing the value of the z variable is shown in Table 5.14 . The same question was modified to collect an arithmetic error from different places in the student workings and to collect other types of errors. Using the analysis' question (questions three), which has been displayed at the end of section 5.5.2, the error data has been analysed.

Table 5.14 Student error collecting question

Student error question
Has an arithmetic error occurred during the computation of the value of z? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank

The result of this analysis is shown in Table 5.15 and Figure 5.30. The table shows 24 occurrences of arithmetic errors. The table also shows the places in the workings that have produced this error and the number of errors in each place.

Table 5.15 Arithmetic error

places_in_workings	num_of_errors	Num_of_errors/24
First element elimination	2	8.33%
Second element elimination	5	20.83%
Third element elimination	5	20.83%
Variable one value computation	2	8.33%
Variable two value computation	6	25.0%
Variable three value computation	4	16.66%
total	24	

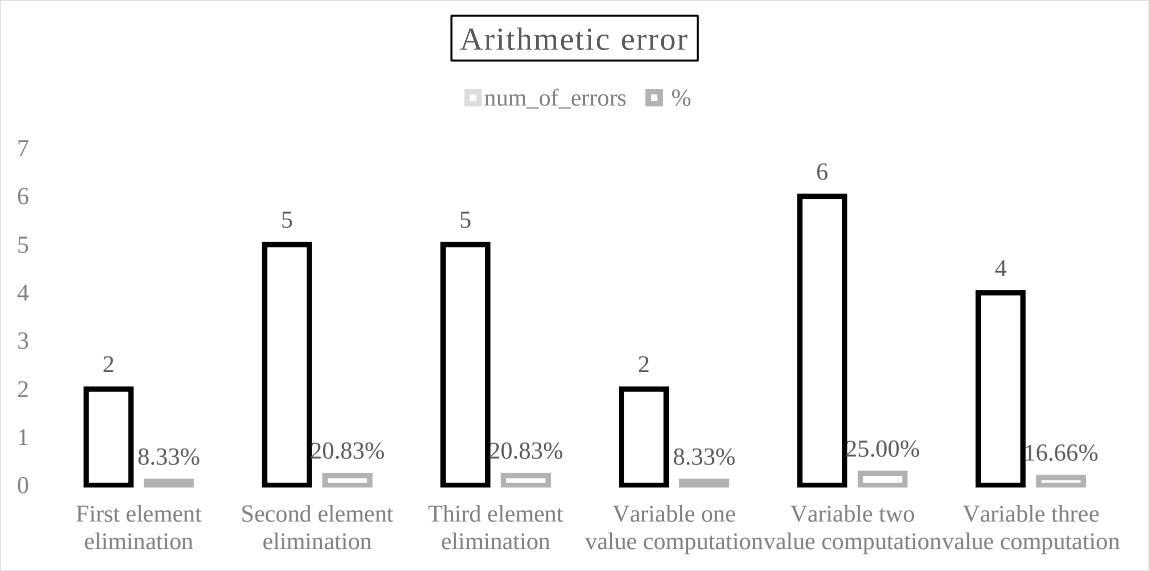


Figure 5.30 Arithmetic error

5.5.2.2 Minus sign error.

Table 5.16 and Figure 5.31 display the result of the analysis of minus sign error.

Table 5.16 Minus sign error data

places_in_workings	num_of_errors	Num_of_errors/23
First element elimination	5	21.74%
Second element elimination	8	34.78%
Third element elimination	7	30.43%
Variable one value computation	2	8.70%
Variable two value computation	0	
Variable three value computation	1	4.35%
Total	23	

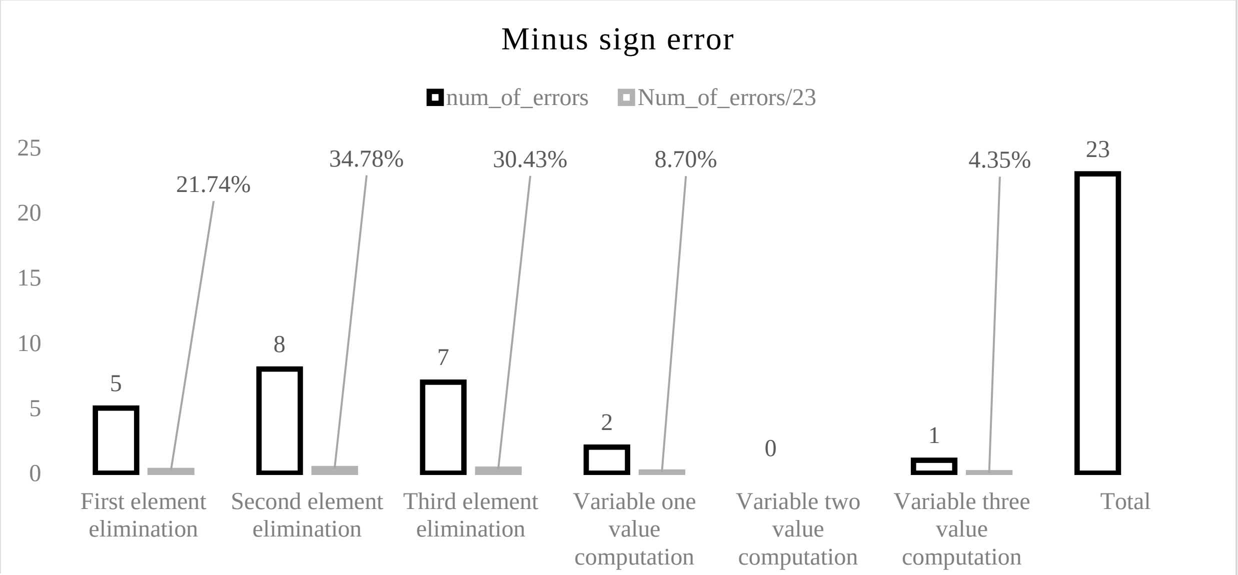


Figure 5.31 Minus error

5.5.2.3 Double minus sign error.

Table 5.17 and Figure 5.32 show the details of the result of the analysis of double minus sign error.

Table 5.17 Double minus sign error data

places_in_workings	num_of_errors	Num_of_errors/24
First element elimination	5	20.83%
Second element elimination	13	54.17%
Third element elimination	5	20.83%
Variable one value computation	1	4.17%
Variable two value computation	0	

places_in_workings	num_of_errors	Num_of_errors/24
Variable three value computation	0	
total	24	

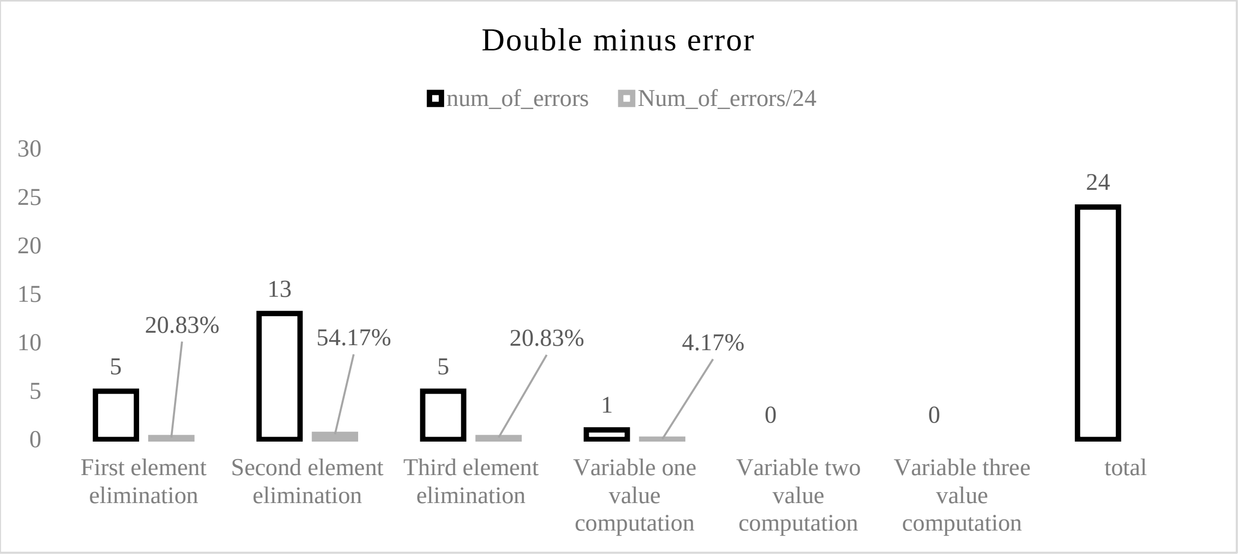


Figure 5.32 Double minus sign error

5.5.2.4 Purposeful error.

Table 5.18 and Figure 5.33 show the result of the analysis of purposeful error type.

Table 5.18 Purposeful error data

places_in_workings	num_of_errors	Num_of_errors/35
First element elimination	6	17.14%
Second element elimination	4	11.43%
Third element elimination	25	71.43%
Variable one value computation	Not applicable	
Variable two value computation	Not applicable	
Variable three value computation	Not applicable	
total	35	

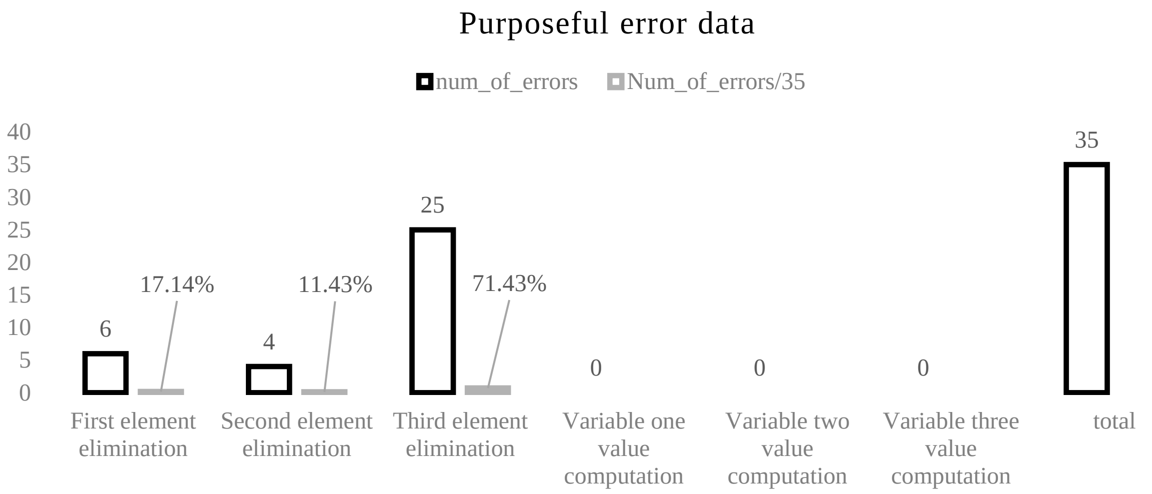


Figure 5.33 Purposeful error data

5.5.2.5 Ignored operation error.

The result of the analysis of ignored error type is shown in Table 5.19 and Figure 5.34.

Table 5.19 Ignored operation error data

places_in_workings	num_of_errors	Num_of_errors/16
First element elimination	5	31.25%
Second element elimination	6	37.5%
Third element elimination	5	31.25%
Variable one value computation	Not applicable	
Variable two value computation	Not applicable	
Variable three value computation	Not applicable	
total	16	

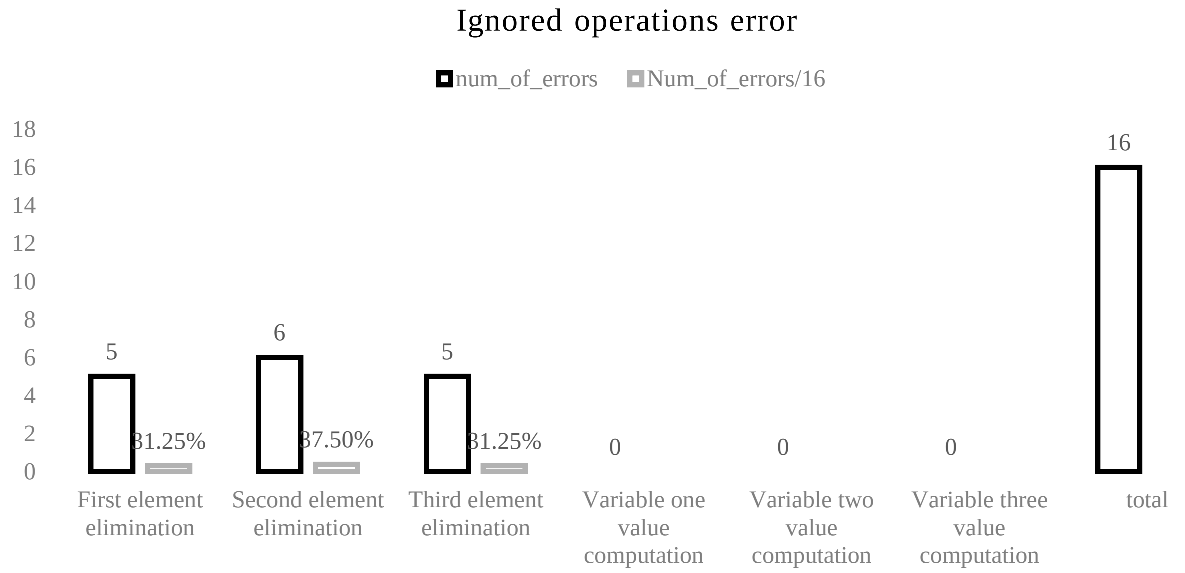


Figure 5.34 Ignored operation error

5.5.2.6 Other errors.

Table 5.20 and Figure 5.35 display the ‘others error’ result of the analysis.

Table 5.20 Others errors data

places_in_workings	num_of_errors	Num_of_errors/27
First element elimination	2	7.41%
Second element elimination	9	33.33%
Third element elimination	16	59.26%
Variable one value computation	0	
Variable two value computation	0	
Variable three value computation	0	
total	27	

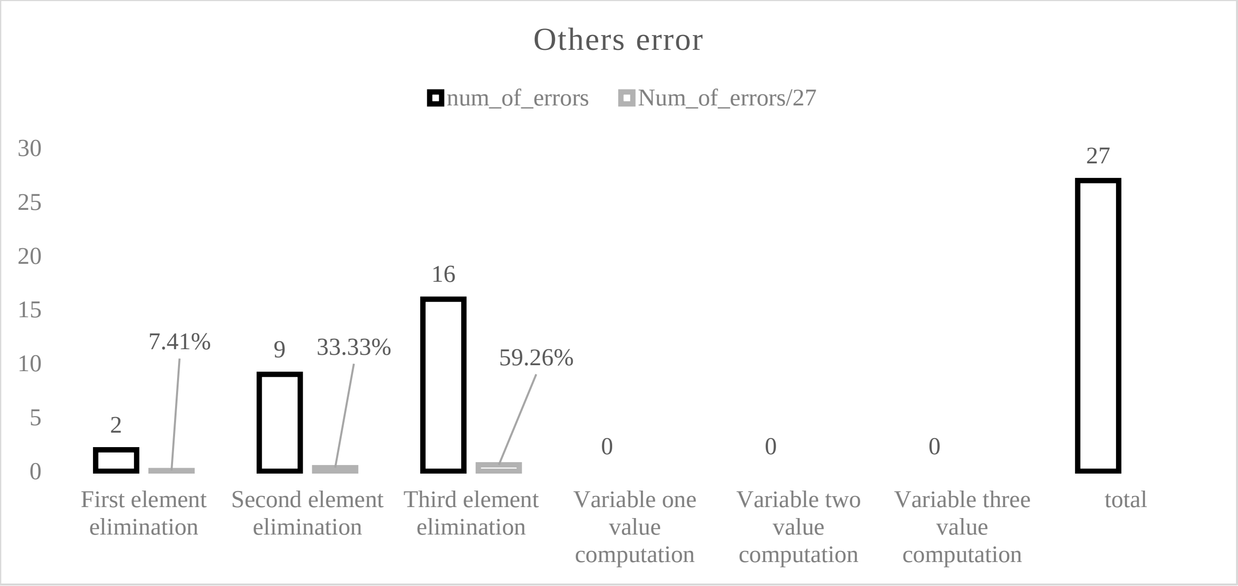


Figure 5.35 Others errors data

5.5.2.7 Solution strategies modification error.

Table 5.21 and Figure *5.36* show the result of the analysis for ‘solution strategies modification’ error.

Table 5.21 Solution strategies modification error data

places_in_workings	num_of_errors	Num_of_errors/8
First element elimination	Not applicable	
Second element elimination	Not applicable	
Third element elimination	Not applicable	
Variable one value computation	2	25%
Variable two value computation	4	50%
Variable three value computation	2	25%
total	8	

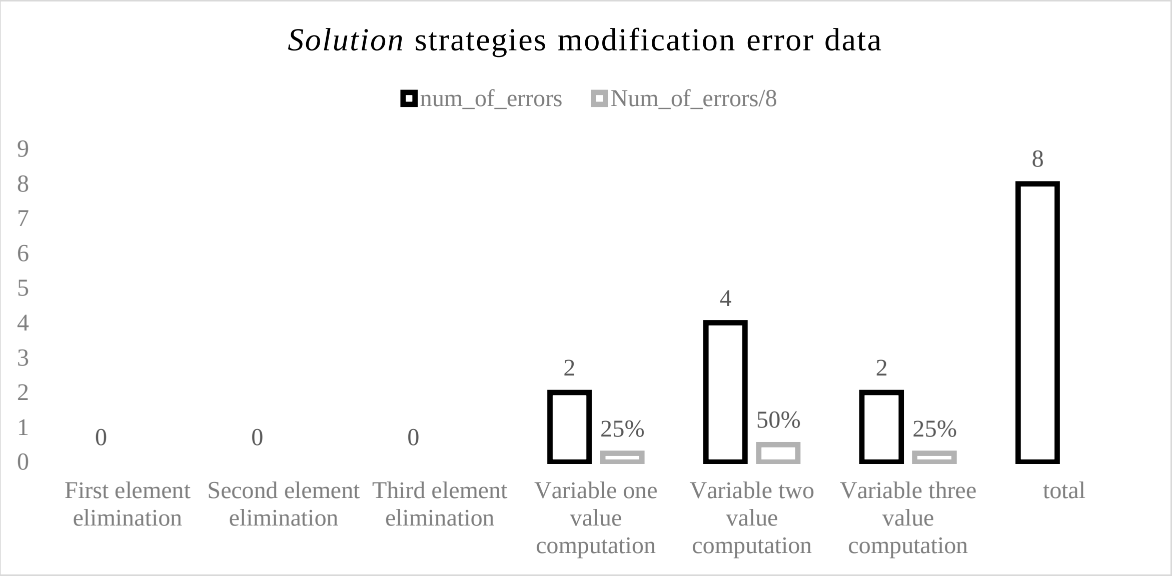


Figure 5.36 Solution strategies modification error data

5.5.2.8 Questions and in-working miscopying error.

Table 5.22 and Figure 5.37 show two sources in the student workings that are responsible for producing the miscopy error type as well as the other details in the result.

Table 5.22 Questions and in-working miscopy error data

places_in_workings	num_of_errors	Num_of_errors/26
Original question miscopying error	22	84.62%
In working miscopy error	4	15.38%
total	26	

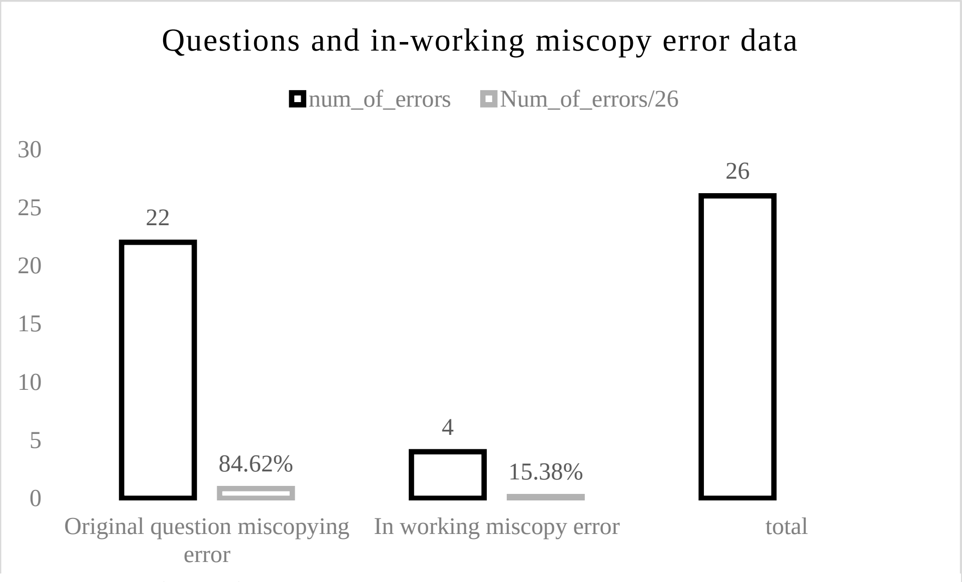


Figure 5.37 Miscopying error

5.5.3 Errors analysis discussion.

This section discusses the analysis of student errors that have been presented in the previous section. The total number of all types of errors in the student workings and the relative frequencies of each type of error are shown in Table 5.23 and Figure 5.38. The total number of errors that have been found in the scripts was 183. The number of errors that have been shown in this table and in the tables in the previous sections do not refer to the number of students who have made any type of error but to the number of errors across all student workings. One student workings might contain no error or any number of errors of one type or multiple types of errors. However, there is one exception, which is the 'original questions miscopy' error. Each 'original questions miscopy' error represents one student.

Table 5.23 Proportions of types of errors in the scripts

Type of errors	num_of_error	Proportions of types of errors = Num_of_error/183
Arithmetic	24	13.12%
Minus sign	23	12.57%
Double minus sign	24	13.12%
Purposeful	35	19.13%
Ignored operation	16	8.74%
Others	27	14.74%
Solution strategies modification	8	4.37%
Miscopying	26	14.21%

Type of errors	num_of_error	Proportions of types of errors = Num_of_error/183
Total error	183	100%

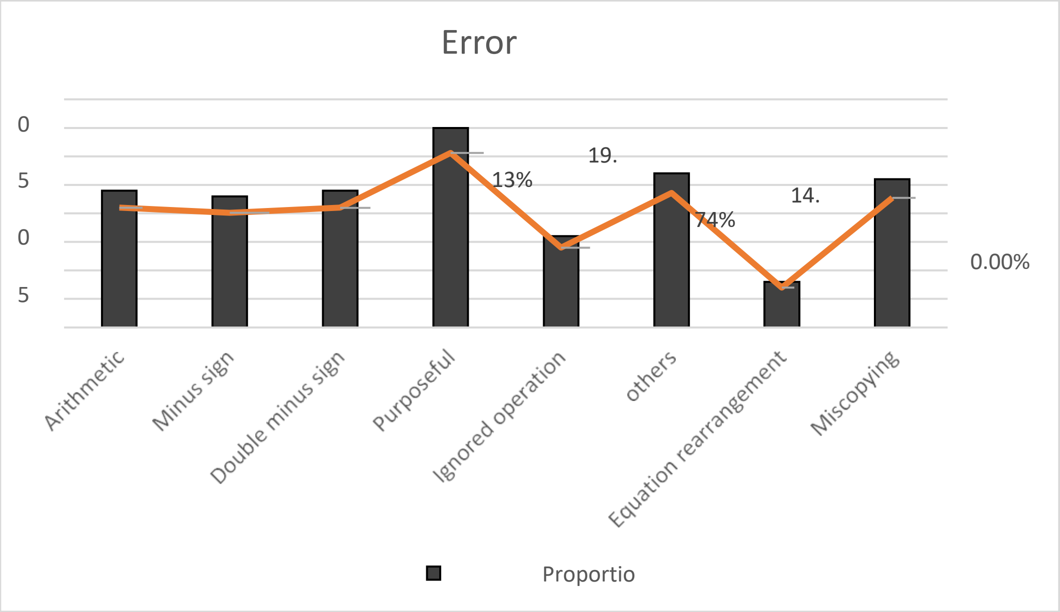


Figure 5.38 Students errors

There are eight types of errors data that have been presented in section 5.5.2. The places (sources) of these errors are displayed next.

1. First element elimination.
2. Second element elimination.
3. Third element elimination.
4. Variable one value computation.
5. Variable two value computation.
6. Variable three value computation.

Some types of these errors appear in all places of the student workings and some only appear in some places. The ‘arithmetic’ (shown in Table 5.15) and ‘in-working miscopy’ (shown in Table 5.22) errors appear in all places of the students workings. On the other hand, the ‘purposeful’ error (shown inTable 5.18), ‘ignored operation’ error (shown in Table 5.19), ‘others’ error (shown inTable 5.20), ‘solution strategies modification’ error (shown in Table

5.21) and 'miscopy ('original questions miscopy' and in-working) errors, (shown in Table 5.22), exist in some places or stages of the student workings. The data referring to these errors including the places in which they appear, have been presented in sections 5.5.2.1 to 5.5.2.8.

In Table 5.23, the most occurring type of error in the student workings was 'purposeful'. The table shows that this error has occurred 35 times. The examination of Table 5.18 (in section 5.5.2.4) shows that the majority (25 times out of 35) occurrences of this error was made when attempting to eliminate the element 8 in the matrix. The elimination of the element 8 is the last task in the element's elimination process. This size of purposeful error might be related to the accumulation of errors in the previous works that involve the elimination of elements 4 and 7.

Table 5.22 also displays that 4.37% of student workings have shown the 'solution strategies modification' error as the least occurring error. Generally, the main cause of this error is the lack of knowledge in changing the sign of a term or terms when manipulating or rearranging an equation. On contrast, the miscopying error is caused because of a human error.

The discovery of an original question miscopy error is very important. The quantitative analysis has shown that 22 students have copied the original questions incorrectly. This number is equivalent to 12% of the total error in the student workings and represents approximately 14% (22/160) of the total students. The decision that assessors make when assessing the student workings that are based on miscopied questions is important. If these workings were excluded because of miscopying of the original question, it may be considered irrational, especially if these students produced correct solutions, as the students have demonstrated their knowledge of the process.

The other error that is not analysed is the ‘follow through’ type. The definition of this error has been given in Table 5.3, section 5.4.1, in the qualitative analysis of the scripts.

All the previous sections in this chapter have covered the presentation and discussion of both qualitative and quantitative phases analyses of scripts. The following section presents the arguments, views and rationales that will be made to claim the outcomes of the analyses of scripts as set of theories.

5.6 Theorising analyses of students working

There are different views in both academic and practitioner communities concerning the acceptance of scientific theories. “For example, some practitioners and academics believe that theory and its application are very limited and, therefore, not very useful in the real world of business. Others feel that very little theory exists in the academic world” (Wacker, 1998, p. 361). However, the presence of “theories vary in their degree of formality or informality. Informal theories are those that are not stated explicitly and do not have … a goal” (Gelso, 2006, p. 3). The informal theories are the theories that are implicit in research. Some researchers accept the presence of a theory in specific disciplines. For example, Gorelick (2011) claims that some researchers associate theory “with mathematics, abstractness, interdisciplinary study, and philosophy” (p.1). Gorelick also states many practitioners have the view of “some gestalt for what does and does not constitute theory” (p.1). In the Oxford dictionary the word ‘gestalt’ means “a set of things, such as a person’s thoughts or experiences that are considered as a single system which is different from the individual thoughts, experiences, etc. within it”. Having such a view might cause confusion because of the separation of an object from its features. The rationale for this confusion about a scientific theory could be because of its widespread definitions. For example, “in her review of the literature, LaJoie (2003) reports finding 20 different definitions of animal-assisted therapy, and 12 different terms for the same phenomenon” (Fine, 2010, p. 34). However, here the focus is only on a few definitions of scientific theory that will achieve the purpose of verifying the knowledge, which has been produced from the analyses of scripts, as theories.

Gorelick (2011) defines theory as the “formation of testable hypotheses (testable so that they fall within the realm of science)” (p. 4). More than one definition is also given of the term ‘hypothesis’: one of these definitions regards a hypothesis as consists of declaration with relations that can be verified; and a hypothesis is used, for example, as the tool to confirm a theory (Vogt, 2005). The definitions of the theory which are given in this section can be compared with the hypothesis definition to understand the differences between both terms. According to Hooks (2000) some feminist studies regard theory as “the development of ideas” (p. 112). Wacker (1998) explains theory in terms of components that include, “(1) definitions of terms or variables, (2) a domain where the theory applies, (3) a set of relationships of variables and (4) specific predictions (factual claims) (Hunt, 1991; Bunge, 1967; Reynolds, 1971)” (p. 363). In this research, the identified and defined terms are created from the discussion and interpretation of the findings from the qualitative analysis of the student scripts. The created theories are applied in the domain of mathematics and in DSR. The knowledge that the theories represent, for example, student errors can be used in the assessment and teaching of the analysed student workings and related topics in mathematics. The knowledge in the theories can be used in creating DSR artefacts that assist in assessing subjective and objective assessments of student workings. The relationships between constructs are also performed when building the DSR model which is similar to the building of a theory. The fourth component of Wacker’s definition which is a specific prediction, can also be performed in the quantitative analysis of the student scripts. Gelso (2006) points that it is “useful to think of any theory as serving certain functions, effectively. Rychlak suggests four such functions: descriptive, delimiting, generative, and integrative” (pp. 3-4). Gelso elaborates on each of the functions. The descriptive function is that “any theory is to describe phenomena” (p. 4). The analyses of student scripts have shown the creation of categories (constructs) that have been used to explain the complexity of assessing student workings. The knowledge inferred from the analyses enables the understanding of students' workings, and the challenges of creating e-assessments to assess subjective questions. Furthermore, it should be clear that "theories constitute a series of related constructs, and vary in their level of abstraction, objectivity, realism, perspective and formality, and must be empirically tested to be called scientific (Gelso, 2006)" (Mouza, 2018, p. 2). The association between science and theory is also discussed by Lee (2014): "a science's description or explanation of what exists or has existed, or how to create what does not now exist or has not existed before, is what I will refer to as 'theory' " (p. 350). The uncovering of the terms which are different types of solution strategies that students applied, and different types of errors in their workings, are regarded as the science of description or explanation. Hence, the categories and subcategories (both solution strategies and student errors) that have been reported in section 5.4 can be claimed as 'good' theories.

For a theory to be accepted as good, it should offer ways to be measured. A theory should enable its empirical testability to be considered a 'good' theory (Wacker, 1998). Gelso (2006) also claims that "a theory ought to tell us *why* the variables or constructs are expected to relate to or influence one another. There must be good reasoning, so to speak, behind the expectation of how variables will be interrelated" (p. 3). In this research, the qualitative data has been quantified to analyse the scripts statistically. The relationships between the constructs have been explained in detail in chapter five and appendix B. On the other hand, some authors separate practice from theory. However, in this research, the analyses of scripts generated theories in the designing and creation of this research's DSR models' artefacts, and this shows the complementary nature of both theory and practice. Wacker asserts that "practice without theory can quickly become a dull and dangerous occupation" (p. 361). The relationships between natural, social and design sciences have been explained in section 3.1. Lee (2014) emphasises that “what is being applied are theories developed by the science in categories i and ii” (p. 351). The categories that Lee is referring to are ‘natural’ and ‘social’ sciences.

In conclusion, the claim that the knowledge from the analyses of scripts has reached the level of scientific theory has been discussed through the identification and discussion of nature of scientific theory, and through reviewing previously published research, in addition to the discussion of the findings in the analysis of scripts in this chapter.

5.7 Chapter summary

The purposes of this chapter were to present the outcome of the qualitative content analysis of student scripts and to present and discuss the quantitative analysis of the scripts. In addition, the chapter performed a review of scientific theories to explore the possibility of claiming the outcomes of the qualitative and quantitative analyses of scripts as a set of theories. Before presenting the outcome of the qualitative analysis, the acceptability of this qualitative data was discussed in terms of reliability, validity and credibility of both the analysis process and its outcome. Examples were given in different sections of this chapter to demonstrate reliability, validity and credibility.

These examples presented and discussed complexities associated with students and assessors’ actions. The discussed complexities in the student workings were (1) understanding these workings so meaningful data can be extracted, (2) difficulty in determining what to collect and what to ignore or leave because some scripts contained contradicting information and (3) difficulty in categorising data that was extracted.

Considerable efforts were made to resolve these challenges. Examples that were discussed regarding assessors’ actions have shown how marks were awarded to students incorrectly. This was demonstrated by showing that assessors had given equal marks to student workings that provided different amount of correct works. Some examples have also shown how assessors awarded the same marks for student workings that showed different amount of correct works. After showing the validity, reliability and credibility of the qualitative content analysis process of scripts and its outcome, the chapter described forming categories and subcategories using findings that came from the analysis.

The sources of names that were used to name subcategories were discussed and names were assigned to categories and subcategories. In total 12 types of subcategories of solution strategy and 11 types of subcategories of students' error were formed. These subcategories and their meanings as well as relationships between them were presented. The knowledge presented in the subcategories and their meanings was used to answer qualitative research questions seeking specific knowledge from the student workings. To demonstrate the validity and reliability of the presented subcategories, examples that showed details of the development of some of the presented subcategories were given.

All the examples that were given in this chapter have used the student raw data, which is student scripts, and the analysis' document that was used to keep records of the process of the qualitative analysis of the student scripts. These two sources have been used to demonstrate challenges in analysing the scripts as well as the outcome of the analysis of the scripts. The next tasks in the chapter was the presentation and discussion of the result of the quantitative analysis of scripts.

At the outset of the presentation and the discussion of quantitative analysis, the explanation of the format that was used to present and discuss the analysis was given. The format included the purpose of analysing data, the questions that collected data to be analysed, the analysis' questions that determine the type of information to be produced as the result of the analysis. And finally presenting and discussing the analysis.

Using this format, two types of data concerning solution strategies were presented and discussed. The first one was to learn from the analysis of solution strategies the orders of operation that students had followed when eliminated elements form a matrix. The data about the number of students that have chosen specific orders was also presented and discussed. The second data that was presented and discussed was about types of solution strategies that were used by students to perform elimination of elements and computation of values of variables. The extent of the usage of these types of strategy was also discussed.

The data in the analysis of solution strategies was used to answer the quantitative research questions that seek knowledge presented in the analysis. The quantitative analysis of the solution strategies has showed that most of the students had followed the expected orders of operation and used the expected types of solution strategies. The other quantitative analysis that was discussed was the student errors.

The same format that was used when presenting and discussing the analysis of solution strategies was used to present and discuss the analysis of the student errors. The outcome of the analysis of eight types of errors that were collected from six different places were presented. The same question was used, or questions were used to collect errors for analysis. To analyse the collected data, the same analysis questions were used to extract the same data from different errors. The outcome of the analysis which is the frequencies of the occurrences of these errors and the processes or the stages that produced them were presented. The details of the analysis of 'arithmetic error' was given as an example. The other types of errors presentations were limited to displaying data that showed only the frequencies of occurrences and their places in the student workings.

The total numbers of all types of errors and each error relative frequencies were given. The analysis of student errors was discussed. The discussion of the analysis of errors showed the existence of some errors, for example, arithmetic error in all six places, while others appeared in some of these places. The importance of collecting the ‘original question miscopy’ error was discussed in relation to the assessors’ ways of assessing student workings. The discussion of the student errors concluded the efforts that have been made to analyse the student scripts qualitatively and quantitatively. The final part of the chapter was evaluating the status of the knowledge in the outcome of the analyses. This evaluation was not performed in terms of validity, reliability and credibility of the outcome, but in terms of theorisation of the outcome.

To achieve this aim, the literature review of scientific theories was performed to understand the nature of a scientific theory. The discussion of different views of scientific theory concluded by selecting the definitions and traits that were relevant to the purposes of the analyses of scripts. These definitions and traits were used to explore all the efforts that have been made to analyse the scripts as well as the outcome of the efforts and the implementation or application of the theory to be built. Based on the outcome of the discussions and exploration, the knowledge in the outcome of the analyses was claimed as a set of theories.

This knowledge will be used, in the next chapter, as foundation to build design science research (DSR) artefacts.

Chapter 6 DSR artefacts of subjective assessment

The previous chapters (two, three, four and five) have contributed to understanding problems and providing knowledge that will be used in the efforts to achieve the purpose of this research. Chapter two (literature review) has uncovered problems of assessing subjective questions of mathematics using e-assessment. To understand the problem further, student scripts have been selected for investigation. Chapter three has investigated the research paradigms and selected design science research as the most relevant paradigm to be used in this project. This paradigm was used to select the most relevant research methods to analyse scripts. Chapter four has shown the usage of these methods to analyse the scripts. In chapter five, the findings in the analyses have been used to produce the knowledge that will be used to create important parts of DSR constructs and models.

The purpose of this chapter is to create selected DSR artefacts using the knowledge that have produced and the methods that have been explained in previous chapters.

Section 6.1 starts with reminding readers about relationships between information systems research and DSR projects that have been covered in sections 3.4 and 3.1. Section 6.1 explains types of artefacts that a DSR produces as well as these artefacts' characteristics. In section 6.2, different definitions of the term 'design' will be given. Furthermore, this section explains types of designs, efforts required to produce relevant designs, and the content of these types of designs. In this section, the definitions of 'design science' (DS) and 'DSR in addition to the relationships between design, DS, DSR and DSR artefacts will also be explained. This section concludes by identifying a starting point of the DSR project regarding these three entities.

Section 6.3.1 begins with defining the term 'construct' in relation to DSR artefacts. The purpose and importance of constructs in a DSR as well as forming relevant constructs will be explained. Furthermore, sub sections 6.3.1.1, 6.3.1.2, 6.3.1.3 and 6.3.1.4 will explain the sources of these constructs as well as presenting the constructs that will be formed.

Section 6.3.2 starts with defining the term ‘model’. An ‘abstract’ or ‘conceptual model’ will be defined in relation to DSR artefacts. This section describes the purpose of models in DSR as the entity that contributes to understanding problems and providing solutions. Steps and sources that will be used to build a DSR model will also be discussed.

A model that shows general utilities of artefacts that implement the method marking concept (MMC) will be presented in section 6.3.2.1. This model shows three main processes (matrix formation, elements elimination and unknowns’ values computation) that student workings contain. Section 6.3.2.2 describes and presents a model that represents scenarios that might happen when students try to form an augmented matrix. Following the formation of this model, section 6.3.2.3 presents models that will be built using constructs that have been generated from the analyses of student scripts. In this section three models will be formed. These models show scenarios that involve the processes of eliminating elements and computing values of unknown. These models display different types of solution strategies that are used during performing the two processes as well as the effects of the presence of errors on awarding marks to student workings. In this section, types of errors that affect scores to be awarded are also listed. The next section deals with the scenarios that entail student interaction with the MMC artefact.

Section 6.3.2.4.1 presents a model that shows activities of students when using MMC artefact to produce their workings and submitting them for assessment. This section defines the main components in this model as well as explaining their tasks.

Section 6.3.2.4.2 defines components of a model that represents assessors’ tasks that include assessing student workings and awarding marks. In this section, the utilities of each component in the model will be explained. Furthermore, this section explains the need of this model for the functions of models that have been created in sections 6.3.2.1 and 6.3.2.2.

Section 6.3.3 defines DSR methods as an artefact that is created from constructs and models. There are multiple techniques to form DSR methods. This section focuses on using the flowchart technique to build this DSR methods artefact. To demonstrate the usage of this technique, flowcharts will be used to show the steps that are required to display an examination question. In addition, this section presents code that will be used to implement the steps in these flowcharts. Finally, this section ends by displaying the results of the execution of the code.

Section 6.3.4 starts with defining an instantiation artefact in terms of DSR. The relationships between this artefact and other DSR artefacts in this research will also be discussed. This section specifies the relationships between models that are developed in this research and the prototype that will implement some part of this model. There are two main parts in this model that address students and assessors' tasks. This prototype will only implement tasks that students perform.

The 'main window' that displays utilities that students can use is explained in section 6.3.4.1. In section 6.3.4.2, the prototype allows students to display a question to form a matrix in a working area and save it. The prototype ensures that students do not submit an empty working. Additionally, this section explains steps that will be taken to save student workings.

The element elimination tasks will be discussed in section 6.3.4.3. The prototype will display a window with main components that allow students to select an element to be eliminated, enter an elimination method, and show their workings in a working area and a button to save their workings. The prototype prevents students to proceed to a working area before providing an element to be eliminated and an elimination method. It also prevents students saving empty workings.

Section 6.3.4.4 will deal with computing values of unknowns. A window that will allow students to select the unknown variable and provide a method that will be used to compute the value of the selected unknown. The components on this window act in similar ways to those components that have been described in section 6.3.4.3. Students can request at any time the prototype to display their workings.

Section 6.3.4.5 will discuss tasks that are triggered by clicking the 'display all workings' button that has been shown in the 'main window'. The prototype gives three options which are 'do nothing', 'modify', or 'delete'. Tasks that will be performed for each option selection will be discussed in this section.

To verify or test the code that is used to perform tasks that have been explained in sections 6.3.4.2, 6.3.4.3, 6.3.4.4 and 6.3.4.5, extra code is added to the prototype. The purpose of this code and the result of its execution will be discussed in each section where these tasks are explained.

6.1 Information systems and design science research artefacts

The discussions that were presented in section 3.4 about research philosophies concerning IS research and design science research have shown the commonalities between them. The definitions of 'information systems' (IS), design science and general meaning of artefacts have also been given in section 3.1 as part of discussing relationships between other sciences and design science. In this research, the intended IS or 'information technology' (IT) or DSR artefacts are specific outputs of design research and are termed 'design research artefacts' (Pierce, 2014). And this type of artefact "often live primarily within the highly specialized contexts of written publications, conference talks and formal exhibitions and demos, along with the conversations, correspondences and thoughts of researchers, designers and other expert practitioners" (Pierce, 2014, p. 736). According to Gregor (2010), IT artefacts, which are the same as DSR artefacts, "concern complex systems" (p. 58). Chatterjee (2010) also emphasises that a DSR project output is required to solve a significant problem by producing a novel artefact. Alternatively, Chatterjee states that the project is required to produce an artefact that significantly enhances the current solution that is used in solving an existing significant problem.

Various types and numbers of artefacts are produced by different DSR research projects. Hevner et al. (2004) and March and Smith (1995) regard DSR outputs as one or more of these products which are (1) constructs, (2) models, (3) algorithms and (4) instantiations. Other authors refer to more types of artefacts besides those that have been mentioned here. In this research the main DSR artefacts are constructs, models, methods and an instantiation that will be produced in a prototype form. Thus, to produce such types of artefacts knowledge of design and design science is important.

6.2 Design science research, design science and design

'Design' is defined in different ways; however, all of the definitions aim is directed towards the same purpose. Walls et al. (1992) define the word design in terms of its meaning as a noun and a verb: As a noun it is a product which is a plan that is intended to achieve something or output something. As a verb it is a process that "plan and proportion the parts of a machine or structure that all requirements will be satisfied" (p. 42). Faste and Faste (2012) and Blessing and Chakrabarti (2009) define design as a means that produces an artefact or artefacts through planning that also includes the communication of the plan. However, Vaishnavi, Kuechler, and Petter (2017) define design as "to invent and bring into being" (p. 3). Vaishnavi, Kuechler, and Petter divide design into two types: routine and innovative. The discussion here is about the latter one. According to Vaishnavi, Kuechler, and Petter, "innovative design may call for the conduct of research (design science research) to fill the knowledge gaps and may result in research publication(s) or patent(s)" (p. 3).

Some authors such as Hevner et al. (2004), Muratovski (2015) and Bayazit (2004) define design as a problem-solving method. Muratovski claims that design can be seen as an evolving concept from problem solving to problem findings, while Bayazit considers design as a "problem solving and decision-making activity" (p.22). Hevner et al. look at design as "a search process to discover an effective solution to a problem" (p. 88). The analysis of the above discussions about design shows that an innovative design is efforts that may require producing knowledge through research or employing existing knowledge or both to produce an innovative and implementable plan, which is fully documented, to solve significant problems or support in solving these problems.

In section 3.1, the definition of 'design science' has been given in the context of its relationships with other sciences. Here the discussion is in relation to the design concept. According to Cross (2006), design science is an approach to a design which involves applying a rationale that is organised unambiguously, and this approach is completely systematic; in other words, design science is the "knowledge designers can use to make their designs" (Aken, 2004, p. 2).

On the other hand, 'design science research' is the "research, aiming to produce valid knowledge for designing" (Aken, 2004, p. 3). Vaishnavi, Kuechler, and Petter (2017) define DSR as the "research that creates this type of missing knowledge using design, analysis, reflection, and abstraction" (p. 4). The 'this' word refers to the design science knowledge such as constructs, models, etc. In conclusion, DSR, 'design science', 'design' and 'design research artefacts' concepts are interrelated. Figure 6.1 shows these relationships in DSR project activities.

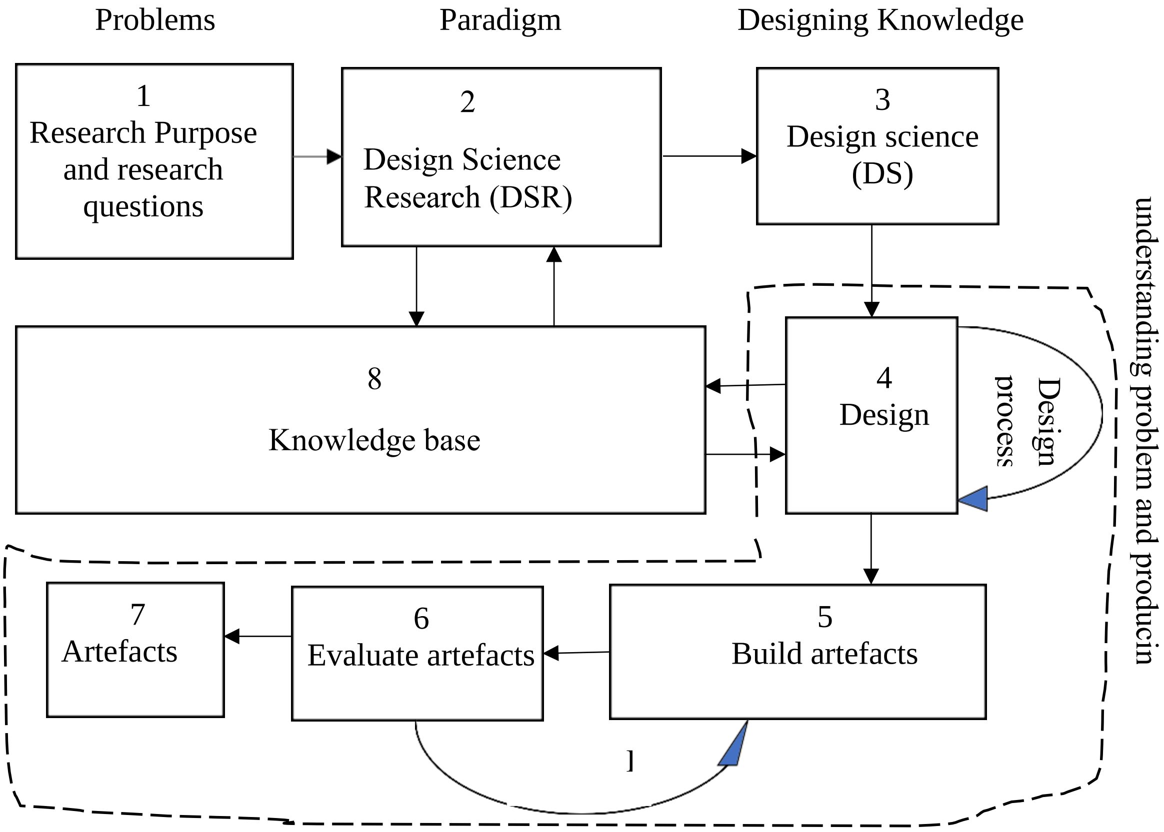


Figure 6.1 DSR, design science, design and design research artefacts

Based on the definitions of these three concepts (design, design science and design science research), the starting point of a DSR project is determined by the nature of the project. Design activities can be the starting point of a DSR project. A design can be considered as a general activity (Faste & Faste, 2012). However, the most relevant starting point of this research project, which is the DSR is labelled 2 in Figure 6.1. The subsequent sections will focus on the creation of selected DSR artefacts.

6.3 Method marking concept (MMC) artefacts

Before building an artefact, there must be a design to build it; because “the process of making something cannot normally start before the process of designing it is complete” (Cross, 2006, p. 3). Cross also suggests another approach to start designing process which is creating an artefact. At this point, there is no design to be followed. Therefore, according to Cross, at this stage, what is required is producing the description of the artefact. This suggestion leads to producing a design or designs using the description of the artefact.

However, Cross' suggestion may be applicable if designers have prior knowledge that satisfies the design requirements. However, providing knowledge of a complex problem requires understanding the problem through investigation. Hence, in this DSR project, the production of knowledge is required to be able to describe that problem and also produce a design. The following sections will focus on producing constructs, models, methods and a prototype using the knowledge that has been produced from the literature review, the analyses of the student workings and other sources.

6.3.1 Method marking concept (MMC) constructs artefact

Constructs are essential elements that are used in defining problems and finding solutions of phenomena (Hevner et al., 2004; March & Smith, 1995; Osterwalder, 2004; , Pereira, Almeida, and da Silva, 2013; Schermann, Böhmann, & Krcmar, 2009; Vom Brocke & Buddendick, 2006; Weber, 2010). They provide "a distinctive medium and language" (Schön, 1983, p. 79) and are considered "concepts that characterize a phenomenon" (Pries-Heje & Baskerville, 2008, p. 733). According to Schermann et al. (2009), constructs "provide the terminology of the research project" (Schermann et al., 2009, p. 185). In DSR, constructs are used to describe purposeful artefacts which might be theory or instantiation or both (Alturki, Gable, & Bandara, 2013). In this research, the complexity of assessing multi-steps questions (MSQ) using e-assessment and the research developed artefacts are described using terminologies that have been created from the analyses of student workings, related domains and activities that have produced these artefacts.

The importance of constructs in DSR has been investigated by Venable (2010). Venable conducted a survey to investigate the views of IS scholars in the DSR field in relation to applying the guidelines given by Hevner et al. (2004). One of the respondents in the survey voiced the view about the significance of constructs in DSR by saying "I don't think that any artefact can be introduced (described) without constructs" (p. 117). Since useful artefacts' creation involves proper design, "part of engaging in good design is choosing a vocabulary or language to use in defining the design task, generating alternatives, and making judgments of balance, fit, and scale" (Collopy & Boland, 2004, p. 14). Great efforts have been made during the analyses of the student scripts and conducting the literature review. These efforts have been discussed comprehensively in previous chapters. The knowledge that has been created from the analyses of student scripts was the source of the main constructs. As for the constructs that come from kernel theories, these are based on the knowledge that have been already accepted and used in their original domains.

Figure 6.2 shows five blocks that are considered the main contributors of constructs. In this section constructs that will be created will be grouped based on their sources and displayed in tables.

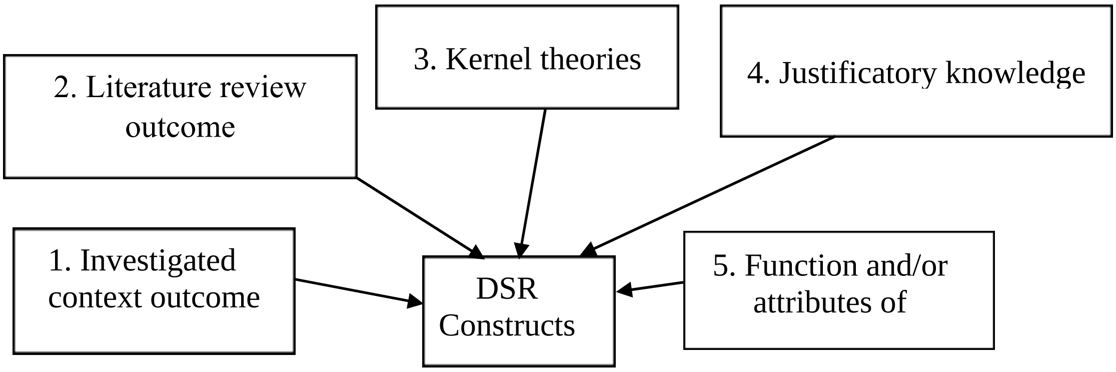


Figure 6.2 Sources of DSR constructs

6.3.1.1 Investigated context outcome.

In Figure 6.2 the first block clockwise is the outcome of the investigation of a context. Using recorded observation and measurement, researchers develop constructs that assist them to explain and visualize the nature of the phenomena and the way they function (Hevner & Chatterjee, 2010). In this research, the literature review and the analyses of student workings are the main means that have been used in understating the context or the nature of the phenomena. The role of the literature review regarding the creation of constructs will be discussed in section 6.3.1.2. In this research, the outcome of the analyses of student workings in examinations is considered the knowledge that is generated from the phenomenon and is used to form constructs. The process of developing categories and subcategories has been discussed in section 5.3. Section 5.4.1 has presented the developed categories and subcategories. The process of developing categories has included among other activities the naming of categories/subcategories of solution strategies and errors as well as their meanings. The purpose of using 'expected', 'acceptable' and 'incorrect' terms for naming solution strategies categories/subcategories have been also discussed in section 5.3. Here, categories and subcategories are all constructs. Therefore, these terms can be used interchangeably. All solution strategies subcategories that have been used to eliminate elements from a matrix and compute values of unknowns are based on these three terms. Therefore, the definitions of these subcategories are not required. Hence, these subcategories names are only given in Table 6.1.

Table 6.1 Solution strategies constructs

Constructs names
Correct expected
Correct acceptable
Correct expected without desired output
Correct acceptable without desired output
Correct expected with extractable info
Correct acceptable with extractable info
Incorrect unacceptable/unrelated with no info
Incorrect unacceptable/unrelated with info
Incorrect unacceptable/unrelated with info and accepting the solution produced
Correct solution strategies that reintroduces
previously solved problems
Excessive/Unnecessary Work with error and/or info

Table 5.1 contains definitions of categories that are used to describe relationships between solution strategies and errors subcategories. The list of these categories' names is given in Table 6.2.

Table 6.2 Solution strategies and student' errors relationships constructs

Constructs names
Desired outputs
Errors
Valid inputs
Awarded

The definitions of all errors' subcategories have been given Table 5.3. Therefore, their names are only displayed in Table 6.3.

Table 6.3 Student errors constructs

Constructs names
Arithmetic error
Minus Sign Error
Double minus Sign Error
Eliminated Variable Only Operation
Follow through error
Ignored/Skipped Operation error
In-working miscopying
Original Question miscopying
Other error
Purposeful error
Solution Strategies Modification

6.3.1.2 Literature review.

The outcome of the literature review also contributes to creating constructs. Regarding the importance of a literature review as one of the sources of constructs, Pereira, Almeida, and da Silva (2013) state that "in order to elicit our constructs and models, a review of prior relevant literature is essential" (p. 5). The findings from the literature review might include knowledge that is based on kernel theories. Therefore, constructs that are extracted from these findings and that are produced in kernel theories might be the same. Therefore, appropriate steps or efforts will be taken to avoid overlapping of constructs. In chapter two, the literature review has identified several subjects/topics that are relevant to the purpose of this research project. These subjects contained concepts and terminologies that are defined or explained; these terminologies and concepts can be accepted as constructs. The subjects that contained these constructs are spread across several sections of chapter two. Table 6.4 displays these constructs and the sections that they appear in.

Table 6.4 Literature review outcome sourced constructs

Constructs names	Definitions	Sources
Academic assessment	A process that involves students performing their tasks and examiners assessing these tasks and giving feedbacks	section 2.3
Objective assessment	Removes examiners effective judgments from assessing students' works while also restricting students' inputs or responses.	section 2.3.2
Subjective assessment	Gives students opportunities to explain their understandings of given questions or works while allowing examiners to make their judgements of student workings.	section 2.3.2
Summative assessment	A form of assessment with aims of measuring current student achievements and awarding grades.	section 2.3.1
Formative assessment	A form of assessment that informs examiners and other stakeholders about current status of students learning levels and help all stakeholders to take appropriate actions to enhance these levels.	section 2.3.1
e-assessment	Tasks that involve using information technology to perform academic assessments.	section 2.4
Ill-defined problems	These are problems with no clear aims causing difficulties in producing solutions.	chapter 2 introduction
Surface learning approach	An approach that shows trivial interest towards learning with the aim of just fulfilling minimum requirements	section 2.7
Deep learning approach	An approach in which students strive to comprehend subject matter they undertake to achieve lasting knowledge	section 2.7
Method marking	A concept that entails awarding marks based on methods that are used in performing tasks rather than on just the final outcome.	chapter 2 introduction
Partial credit (partial mark)	It is a mark assigned to a correct working in a task that also includes incorrect workings	chapter 2 introduction
Follow through error	This involves when student workings that contain error loses mark only once and the effect of this error is not transmitted into another part of the workings.	chapter 2 introduction

6.3.1.3 Kernel theories.

The artefacts that are needed to perform tasks or solve problems reside in contexts; these contexts have associated disciplines or fields. These disciplines provide knowledge (kernel theories) that are required in contexts. According to Walls et al. (1992), kernel theories refer to theories from natural science, social science and mathematics, and they are utilized in design theory. In this research project, the knowledge in kernel theories is specifically present in the Gaussian system of linear equation and mathematics generally. In section 6.3.1.2, relationships between kernel theories and the outcome of literature review have been explained. Based on the relationships explained, the source of kernel theories and the outcome of the literature review could be regarded as one source. However, the outcome of the literature review covers more subjects than kernel theories that are specific to particular parts of this research. Therefore, for clarity the two sources are discussed in separate sections and the constructs that are made from them are also presented separately. Table 6.5 shows constructs that come from kernel theories with the definition of each construct.

Table 6.5 Kernel theories sourced constructs

Constructs names	Definitions
Back substitution	The process that involves substituting a value of known variable in an equation to compute the value of unknown in an equation.
Element	A number/coefficient value that is used to form the augmented matrix.
Gaussian elimination	Solving a matrix equation by using row operation to transform the augmented matrix into upper triangular form (Jain & Gunawardena, 2004).
augmented matrix	"An augmented matrix for a system of equations is a matrix of numbers in which each row represents the constants from one equation (both the coefficients and the constant on the other side of the equal sign) and each column represents all the coefficients for a single variable" (Algebra - Augmented Matrices - Pauls Online Math Notes)
Row operation	Row operations involve manipulating systems of linear equation to find solutions of the system (Solving Systems of Linear Equations; Row Reduction - HMC Calculus).
Size	Number of rows and columns

Constructs names	Definitions
Upper triangular matrix	“An upper triangular matrix is a square matrix in which all entries below the main diagonal are zero (only nonzero entries are found above the main diagonal - in the upper triangle)” (Diagonal Matrices Upper and Lower Triangular Matrices - Faculty).

6.3.1.4 Justificatory knowledge and function and attributes of artefacts.

This section explains justificatory knowledge, and functions and attributes of developed artefacts as being the sources of constructs. According to Gregor and Jones (2007), justificatory knowledge has a wider meaning than kernel theory; it includes all information that informs and enlightens design research. This information could be knowledge from research fields, informal knowledge and the experiences of practitioners. Hence, “relevant constructs are likely to be drawn from justificatory knowledge” (Offermann, Blom, Levina, & Bub, 2010, p. 299). The explanation about justificatory knowledge shows that sources of constructs are not limited. However, in this section only two sources will be used to form constructs that will be displayed in Table 6.6. One of these sources is memos that were used to capture this researcher reflections regarding subjects under investigation and activities undertaken. The second source is functions and attributes of models and instantiation artefacts that have been developed in this project.

Table 6.6 Justificatory knowledge sourced constructs

Constructs names	Definitions	Sources
Actors	“Actors represent the role of the future users of the system” (Papajorgji & Pardalos, 2014, p. 47)	
Assessors	Assessors are one of the actors. They assess student workings based on criteria, assign scores to works and give feedback.	
Assessors’ tasks controller (ATC)	ATC is a component of ’assessors tasks scenarios model’ that allows assessors to retrieve student workings and the associated questions.	Section 6.2.3.4.2
Input and output tools (IOT)	IOT are components (tools) in ‘student tasks scenarios model’ and they assist students to perform their tasks.	Section 6.2.3.4.1

Constructs names	Definitions	Sources
marks	A judgment, expressed as a number or letter, about the quality of a piece of work done at school, college or university **(**Cambridge Dictionary online**)**	
Questions storage (QS)	QS is the same component that is shown in both assessors' tasks scenarios model and student tasks scenarios model; and it is used to store questions.	Section 6.2.3.4.1
Solution type (ST)	ST is part of student workings and it is one of the components in assessors' tasks scenarios	Section 6.2.3.4.2
Students	Students are one of the actors who request examination, record their workings and submit them for assessment.	
Students tasks controller (STC)	STC is a component in the student tasks scenarios model and it allows students to request for questions and working area.	Section 6.2.3.4.1
Subsequent step (SS)	Subsequent step is the next step in student workings. It is independent of previous and next steps workings although it might use an input from previous step and produces output.	
Working area (WA)	WA is used by students to record their workings. It is one of the components of the student tasks scenarios model.	Section 6.2.3.4.1
Working tools (WT)	WT is a component in the student tasks scenarios model, and it allows students to request WA and IOT components	Section 6.2.3.4.1
Workings	Workings are the student's works that show details of their workings including strategies they use to produce answers at different stages.	

Table 6.6 contains dynamic constructs that expand as the functionalities and attributes of artefacts changes. The number of constructs that are required to form models or explain a phenomenon should not be predefined limit. However, the main constructs that are needed to achieve purposes of research must be formed. Constructs creation process is iterative because it is part of a build and evaluate process that continues until a conclusion is made. Therefore, construct creation will continue till the completion of the development of the last DSR artefact in this project.

In section 6.3.2, the constructs that are presented in this section will be used to form models.

6.3.2 Method marking concept (MMC) model artefact.

The term 'model' is defined differently in different fields. Venkatesan and Kuppuswami (2008) categorise models as 'concrete' and 'abstract'; an abstract model is also known as a 'reference model' or a 'conceptual model'. In this research the focus is on the abstract model type, which is a DSR model artefact. In this discussion the term conceptual model and reference model are used interchangeably. The definition of model in DSR project is "a set of propositions or statements expressing relationships among constructs" (March & Smith, 1995, p. 256). March and Smith (1995), Venkatesan and Kuppuswami (2008), Venable (2006) and Baskerville, Baiyere, Gregor, Hevner, and Rossi (2018) define conceptual model as the representation of relationships between constructs or entities; and "in design activities, models represent situations as problem and solution statements" (March & Smith, 1995, p. 256). The analyses of scripts have shown usage of different types of solution strategies to eliminate elements and compute values of unknowns. In this research, examples of models representing relationships between constructs will be shown in sections 6.3.2.2, 6.3.2.3 and 6.2.3.4.

According to Hevner et al. (2004), models help in comprehending problems and finding solutions as well as linking parts in problems and solutions. An example of a model functioning as Hevner et al. suggest has been demonstrated in Table 5.2 and the details of the functions of models will also be shown in section 6.3.2.3 and section 6.3.2.4 and its subsections.

Conceptual models are used in "design problem and solution spaces" (Baskerville et al., 2018, p. 362) and in representing "the intended software in abstract form" (Venkatesan & Kuppuswami, 2008, p. 75). Generally, a conceptual model "forms the foundation for developing the computer code" (Robinson, 2015, p. 5) and they are viewed as "(situated) utility" (Vaishnavi, Kuechler, & Petter, 2017, p. 14). In this research, models that will be created will be used as foundation for creating 'methods' and instantiation. These models will display tasks to be performed and methods of performing tasks.

The formation of models is achieved by analysing relationships between constructs (Weber, 2010). Mwilu, Comyn-Wattiau, and Prat (2016) also specify models as well-defined bodies that are built from constructs. Batani et al. and Bodart et al. (as cited in Vom Brocke & Buddendick, 2006), Jeusfeld, Jarke, Nissen, and Staudt (1998), Venkatesan and Kuppuswami (2008), Shanks, Tansley, and Weber (2003) and Wand and Weber (2002) state that models are generally presented in graphical from. The sources of constructs have been identified and explained in section 6.3.1

Constructs that have been formed in section 6.3.1 will be used to build models. When forming models, firstly relationships between constructs will be determined and secondly, purposes and/or functions of each model will be given. When providing purposes and functions of models that are based on student working, only brief explanations will be given because the required details of these models have already been given in Table 5.2. However, as for constructs that are not in the analyses of scripts, any new model that is formed from these constructs will be explained fully. For clarity, separate models will be created to show different sections of student workings demonstrating how knowledge that has been extracted from the analyses of student workings is used to understand their workings and assist in producing solutions.

In this research, when constructing models, two types of shapes and one-direction arrows are used. The rectangular shapes are used to represent processes and objects.

Processes in these models perform tasks while objects might represent outputs of these processes or inputs required by these processes or components of other objects. The oval or round shapes are used to show attributes of objects. The arrows are used to show relationships between processes and/or between processes and objects. Messages or labels are attached to these arrows to convey specific tasks to be performed or to show objects owning attributes. Rectangular shapes are also used to represent users of the system that these models present and describe.

6.3.2.1 MMC general student workings model.

Figure 6.3 shows a model with three main activities or processes as well as outputs from these processes. The processes that are shown in this model are matrix formation, elements elimination and variables values computation; and the outputs that are produced by these processes are matrix, upper triangular matrix and values of unknown objects.

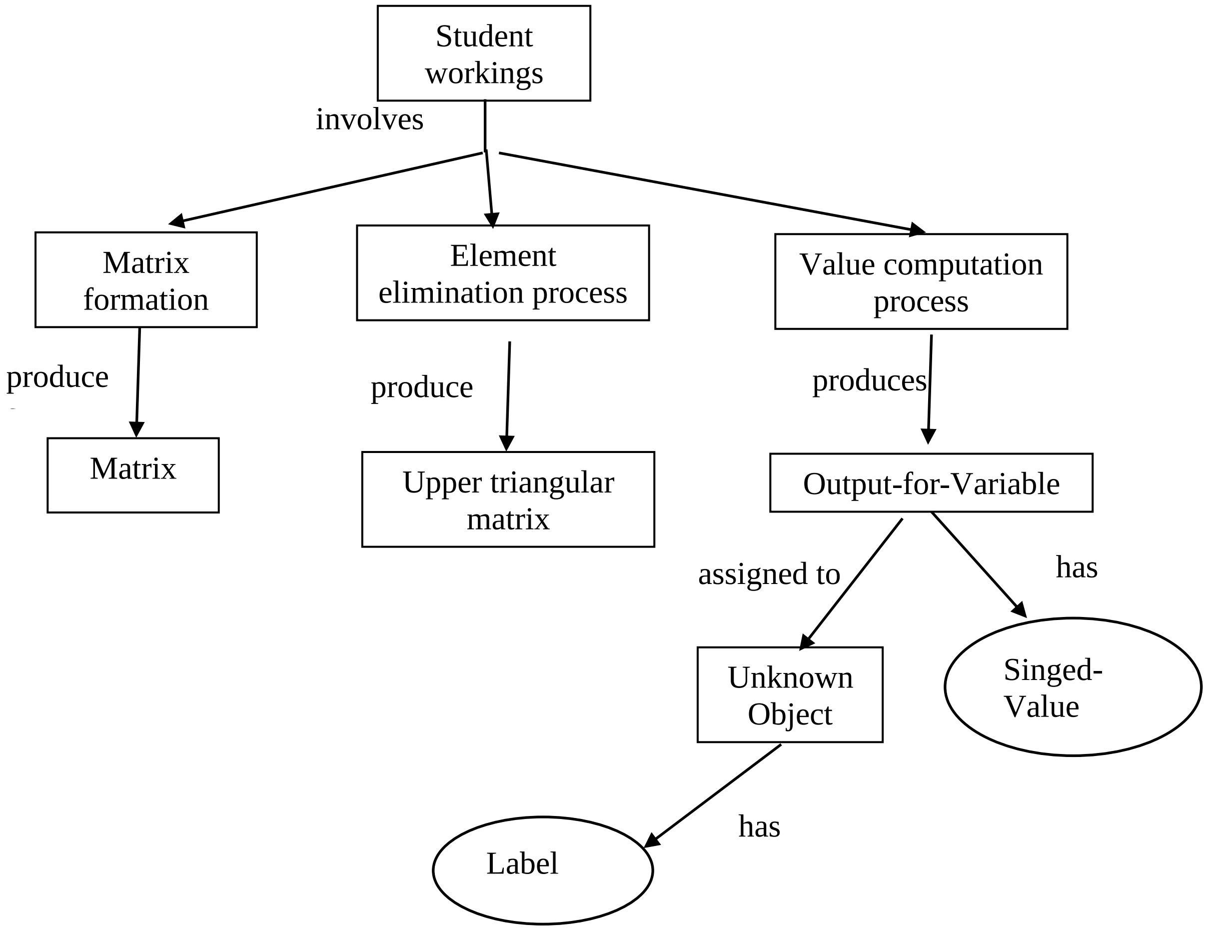


Figure 6.3 MMC general student workings model

The models that contain the details of these processes and their outputs will be given in the following sections.

6.3.2.2 Matrix Formation process model.

The process of forming a matrix of coefficients of the system of equations involves writing equations as a single augmented matrix. Coefficients of variables and values of equations are arranged in two dimensions having rows and columns. Figure 6.4 shows a possible model describing students' efforts in creating a matrix. If students perform a correct transformation of equations into matrix, the size of matrix is given by the numbers of rows and columns. The number of elements in the matrix is equivalent to the product of number of rows and columns. Elements are numbers that are used to form a matrix. The element could be referenced using a row and a column or using its position in the matrix. When student workings produce a correct matrix, they are awarded full marks that is assigned for this part of the workings.

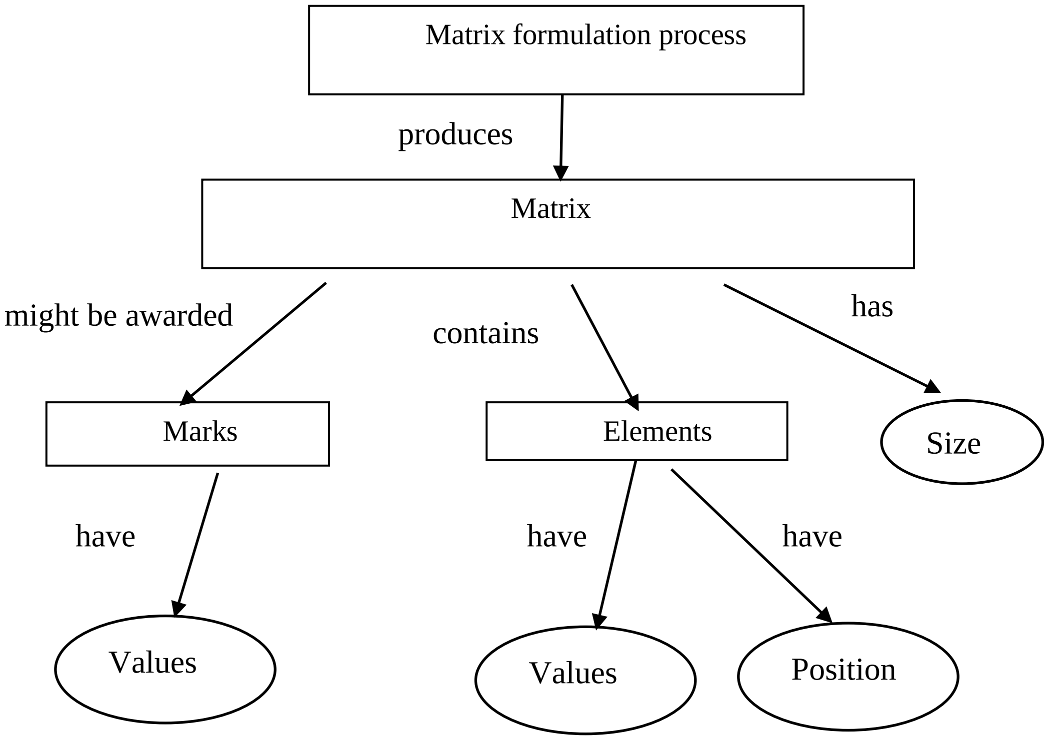


Figure 6.4 Matrix formation process model

However, students might produce an incorrect matrix that might have an incorrect size (producing less or a greater number of rows and/or columns in the matrix), or incorrect elements values which also include incorrect sign. Consequently, such student workings are assigned partial or zero marks. The matrix formation process does not include any computation. The main error that related to the formation of a matrix was the 'original question miscopy error' or lack of knowledge to form a matrix. Here the decision to continue the assessment of the student workings is dependent upon the assessor. This decision will also affect the awarding of marks for this section and the subsequent student workings. Provided the matrix is formed correctly or the decision to continue assessing the student workings regardless of the miscopying the original question is made, the following task in the student workings is eliminating elements from the matrix and computing values of unknown objects.

6.3.2.3 Element elimination and value computation processes models.

The analyses of scripts have shown usage of different types of solution strategies to eliminate elements and compute values of unknowns. Table 5.1 contains 'desired output', 'Errors', 'valid input' and 'awarded' constructs and their definitions. Table 5.3 contains all subcategories of 'errors' constructs and their definitions. Table 5.2 shows each solution strategy definition, function and relationships with other constructs in Table 5.1 and Table. Hence, the content of Table 5.2 is considered a model in a written format. The components of this model will be reanalysed and divided into related groups. These groups will be presented as models in their own right in graphical format. The purposes of dividing this model into sub models are two folds. Firstly, here multiple models are drawn to show details of the model visually and clearly. Secondly, redrawing them in this format enables linking them to other models that need different services of these models.

To avoid a complexity of these models related solution strategies will be grouped and shown in separate models. The differences between all types of solution strategies have already been shown in Table 5.2. The model that shows the usage of 'correct expected' and ‘correct acceptable’ solution strategies includes strategy types 1, 2, 3, 4, 5 and 6. Solution strategy type 10 also uses ‘correct expected’ and/or ‘correct acceptable’ strategies. However, this solution strategy type is different than types 1, 2, 3, 4, 5 and 6. Therefore, it is represented in a separate model. The other model groups solution strategies types 7, 8, 9, 11 and 12. Figure 6.5 shows a model that presents student workings, which use solution strategy types 1, 2, 3, 4, 5 and 6, when eliminating elements or computing values of unknowns. Full explanations and functions of these solution strategies are given in Table 5.2 as well as Table B.1, Table B.2 and Table B.3.

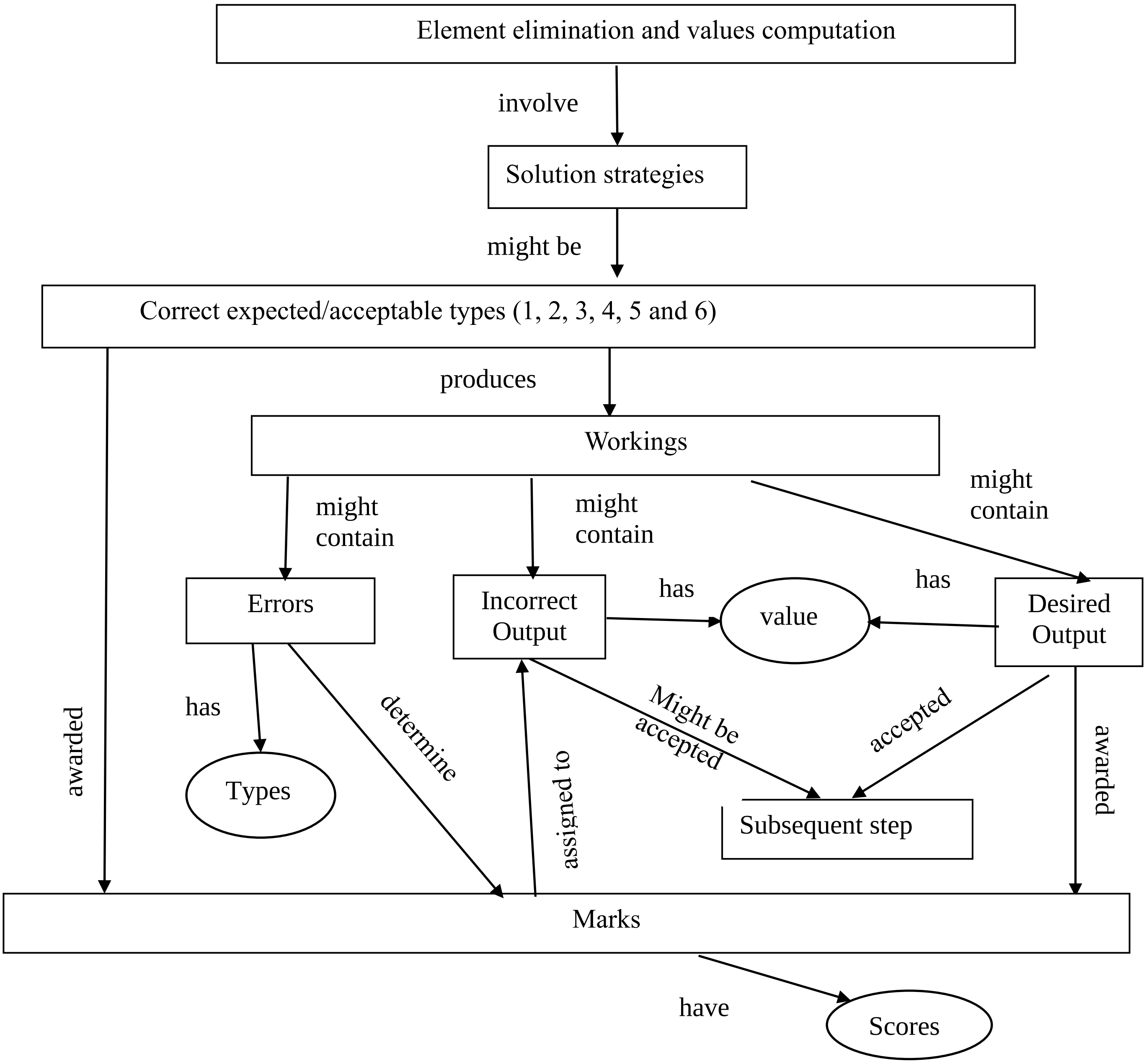


Figure 6.5 Correct expected and acceptable solution strategies application model

Another version of correct solutions strategy (type 10) task and the consequences of its usage is shown in Figure 6.6. The detail about this strategy (model) has been given in Table 5.1 and Table 5.2 as well as Table B.7.

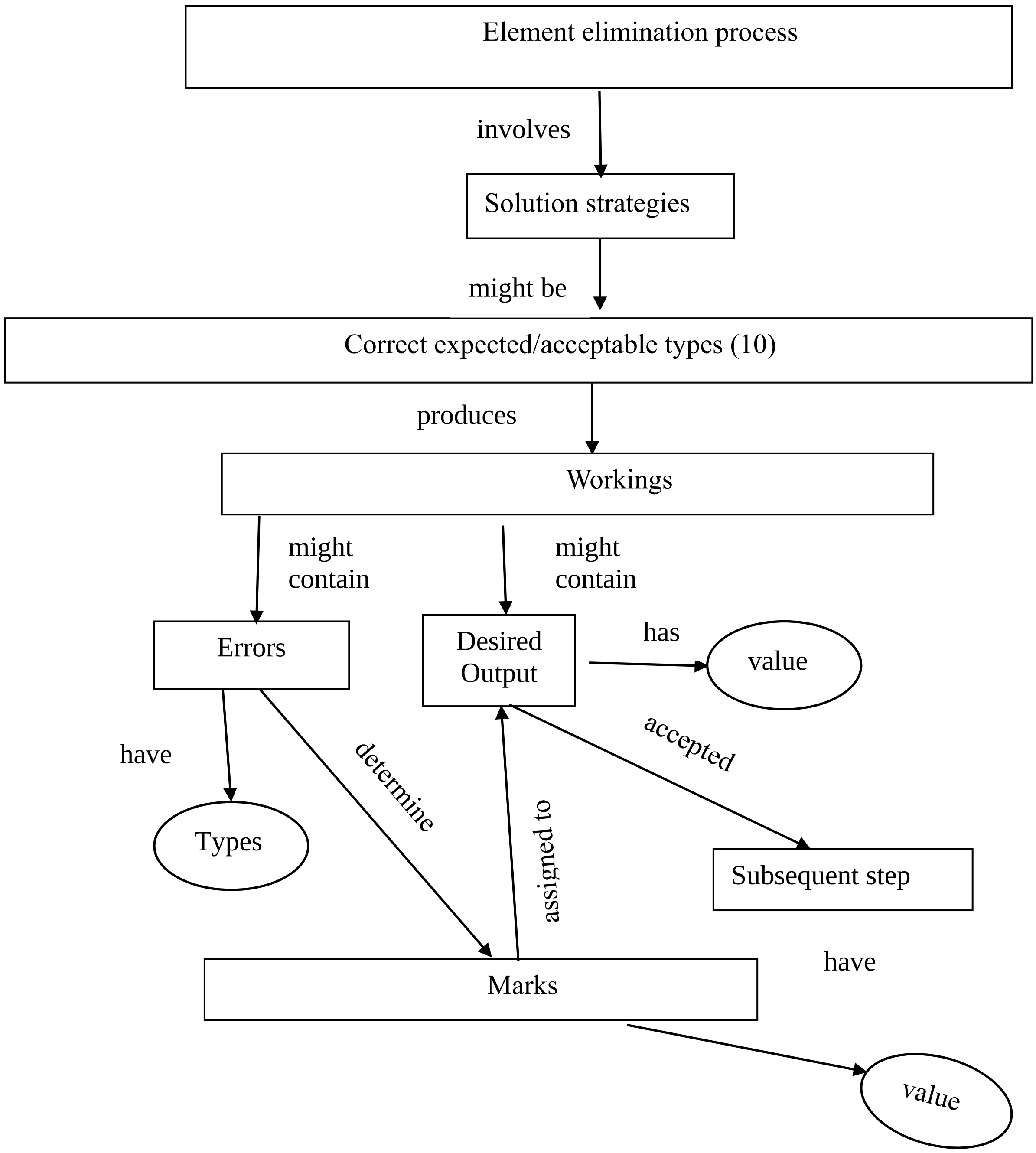


Figure 6.6 Correct expected and acceptable solution strategies (type 10) model

In chapter five, Table 5.3 contains only definitions of all subcategories of ‘Error’ category. Table B.9, in Appendix B, displays the following information concerning these errors.

- ‘Definition’: definition of this subcategory of error.
- ‘Source’: where this error is originated.
- Cause of error’: tasks that cause this error.

- ‘Consequence of error on performing other tasks’: decision of allowing the output of the current step that produced this error as an input in the subsequent task operation.
- ‘Consequence of error on awarding mark’: decision of whether the workings that contain this error is awarded a partial mark or not.
- ‘Comment’: decision to check for this error when assessing student workings.

The information in these tables is necessary for developing and evaluating instantiation part or all the components in models that have been shown above.

The error types in Figure 6.5 and Figure 6.6 show the possibility of student workings containing any errors in Table 5.3. The analyses of student workings have shown types of errors that assessors consider awarding partial marks. These types of errors are listed below. The list shows the name or caption of the table, in Appendix B, preceded by the name of the subcategory’s name. This table contains the abovementioned information about this error.

1. Arithmetic (information in Table B.9).
2. Ignored operation (information in Table B.16).
3. Minus sign (information in Table B.14).
4. Double minus sign (information in Table B.15).
5. Eliminated variable only (information in Table B.17).
6. Purposeful (information in Table B.12).
7. Others (information in Table B.13).
8. Solution strategies modification (information in Table B.11).
9. In-working miscopy (information in Table B.19).

The models that have been presented in Figure 6.5 and Figure 6.6 show that errors effect the amount of marks that are assigned to student workings.

The model that is displayed in Figure 6.7 shows the solution strategies types 7, 8, 9, 11 and 12. Further details of these strategies have been given in Table 5.2, Table B.4, Table B.5, Table B.6 and Table B.8. This model shows that student workings with these types of solution strategies are not awarded any marks. However, as explained in the tables mentioned above, the incorrect output that is produced using strategy type 9 is accepted as a valid input in the next step operation.

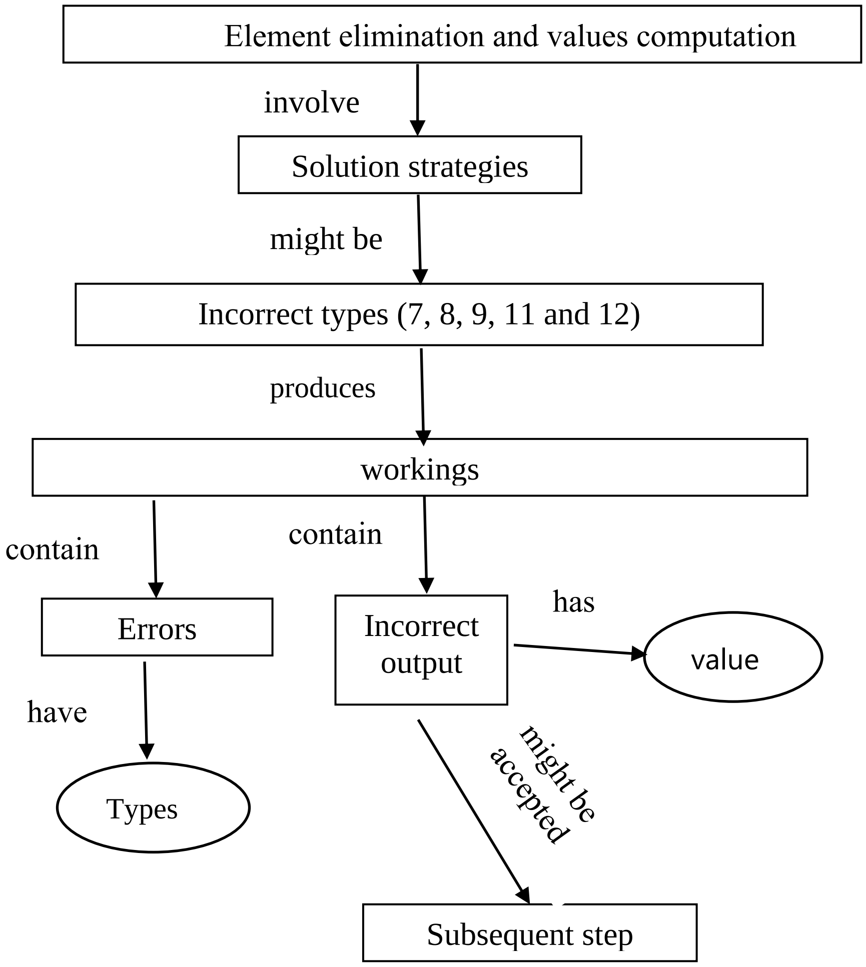


Figure 6.7 Incorrect solution strategies model

In the following sections, models that use services of the models that have been shown in Figure 6.3, Figure 6.4, Figure 6.5 and Figure 6.7 will be built and presented.

6.3.2.4 Students and assessors' tasks scenarios models.

Students in an examination context prepare their workings and submit them. Student workings are kept separately in a safe area, so in future, if needed, they can be retrieved for any legitimate purpose. On the other hand, examiners assess student workings, award marks and give feedback to students. The details of assessment also should be kept separately for integrity or future retrieval in case dispute arises. These are normal scenarios; however, performing these tasks using e-assessment tools for some types of questions, such as subjective questions, is challenging.

6.3.2.4.1 Students tasks scenario model.

Figure 6.8 shows a model that contains two blocks, among others, which are labelled students and actors. However, the focus of this model is about activities that students perform. The assessors' activities will be shown later in section 6.2.3.4.2. This model shows different blocks that represent student's activities. The meanings and the purposes of important blocks are explained next. In this model, four rectangular blocks are numbered 1 to 4. Students' tasks controller (STC), block 1, is the interface object that links students to tasks that they perform. Question storage (QS), block 2, is a storage object that stores questions to be solved. Working tools (WT), block 3, is also an interface object that contains working area (WA), and input and output (IOT) objects. WA is a blank space that allows students to type in their inputs which represent their answers (workings). IOT includes input and output tools such as drop-down list, text field, buttons and labels. These tools allow students to perform different actions such as selecting tasks to be performed, responding to the system's request and submitting their workings. Student workings storage (SWS), block 4, is a storage object that might be files or database tables used to store student workings.

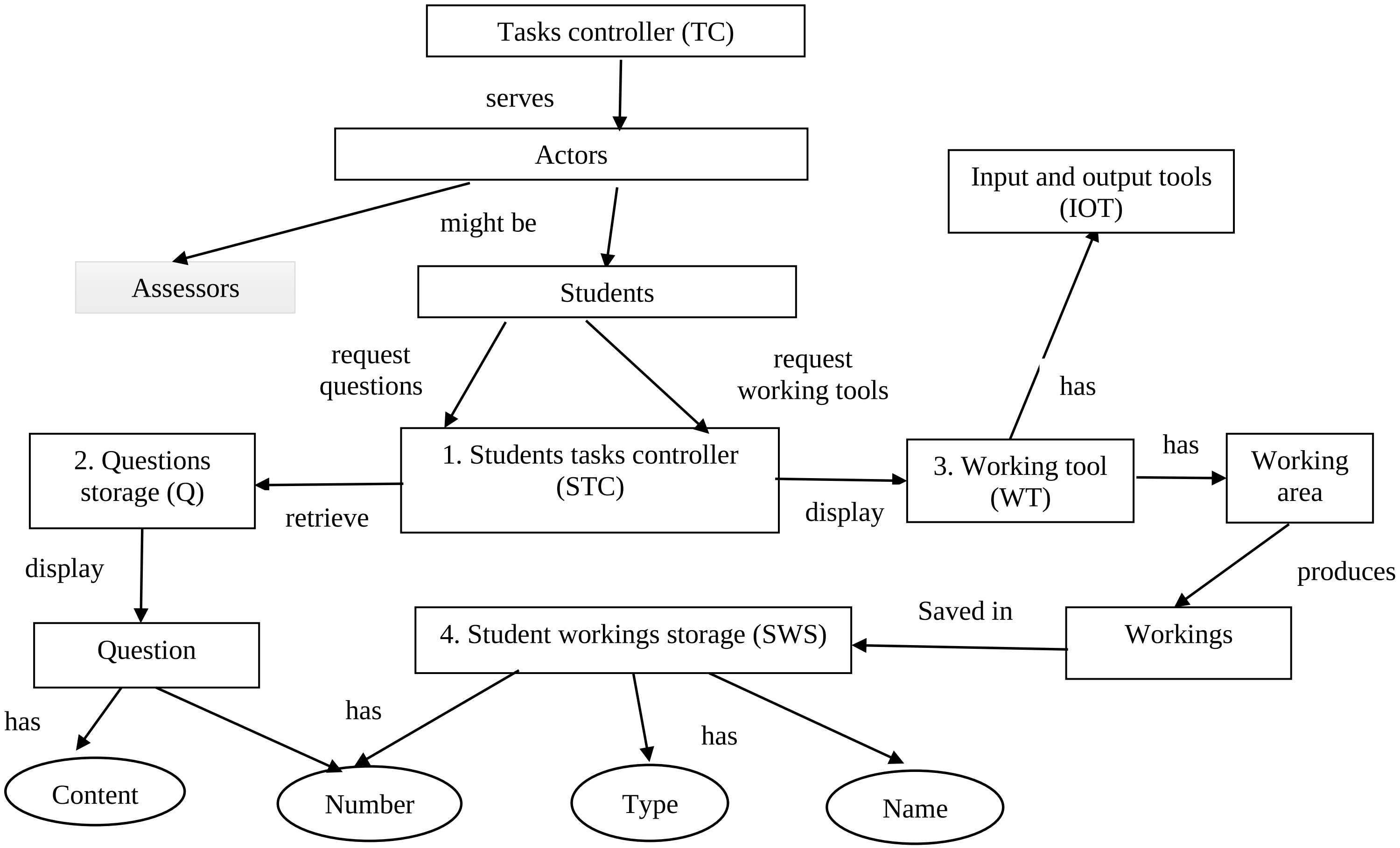


Figure 6.8 Student tasks scenarios model

Student's activities start with requesting questions and WT from STC. The STC block retrieves the requested questions from QS and displays them. It also displays WT. Students use WT to record their workings and save them in SWS. The SWS saves student workings using details, which are types of solutions (type of workings), identification numbers (working number associated to question number) and location name (file). The student workings that have been saved in SWS are assessed by assessors.

6.3.2.4.2 Assessors tasks scenario model.

The assessors' activities model is shown in Figure 6.9. In this model, there are seven rectangular blocks that are numbered. Assessor tasks controller (ATC), block 1, has similar action as the STC block. It is the interface object that connects assessors with SWS and QS objects. SWS and QS are the same components that have already been defined in section 6.2.3.4.1. Solution type (ST), block 4, is a solution type in student workings that is stored in SWS. Matrix formation (MF), block 5, is the type of student workings which show matrix formation. Element elimination (EE), block 6, is the type of student workings that involve the elimination of an element. Unknown object value computation (UOVC), block 7, is another type of student workings that belong to the computation of a value of unknown.

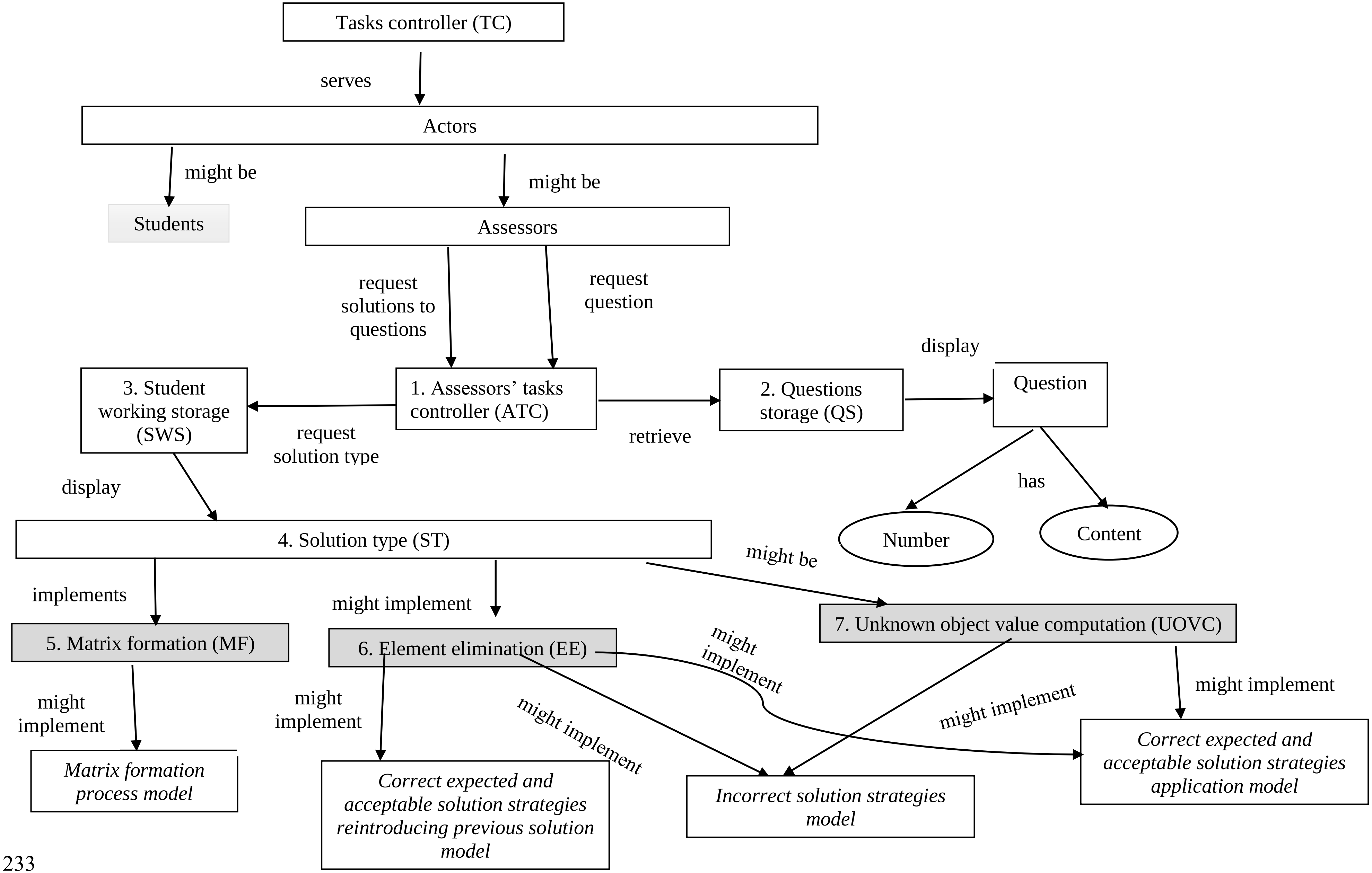


Figure 6.9 Assessors' tasks scenarios model

The starting point, in this model, is when the assessor requests question and the student workings. The ATC responds with retrieving the question from QS and displaying it. The ATC also requests the ST that have been saved in SWS. The SWS displays the ST that might be one or more of the three types that are represented by blocks 5, 6 and 7. Based on the ST retrieved from the SWS, the assessor implements relevant actions which could be any of those that are shown in blocks 5, 6, and 7 in Figure 6.9. If the retrieved ST is MF, the function to be implemented is present in the 'matrix formation model' which is displayed in Figure 6.4 If the ST is EE, the function to be implemented is in one of the models that are displayed in Figure 6.5, Figure 6.6 and Figure 6.7. If the ST is UOVC, the assessor chooses one of the models that are displayed in Figure 6.5 or Figure 6.7 to implement the required function.

In this research, the models that have been built collectively form a DSR model. Selected parts of this model will be used to create a prototype that will be explained in section 6.3.4.

6.3.3 Method marking concept (MMC) methods.

This section discusses the development of DSR methods and demonstrates the implementation (instantiation) of these methods. The creation of a prototype of a DSR instantiation requires DSR methods. These methods are algorithms that are used in performing a specific task (Brandtner, 2017; Vaishnavi, Kuechler, & Petter, 2017). Methods can also be applied to more than one problem or task. Methods can be "used to solve a specific problem or classes of problems (e.g., DSR method for conducting DS research, algorithms)" (Alismail, Zhang, & Chatterjee, 2017, p. 23). They are also seen as "instructions for performing goal-driven activities" (Gregor & Hevner, 2013, p. A3). The construction of methods is based on constructs and models (Brandtner, 2017). Thus, in this research project constructs and models that have been presented in sections 6.3.1 and 6.3.2 will be the sources of methods that will be developed to implement some parts of this artefact.

There are numbers of techniques to create DSR methods artefact. Pseudocode and flowcharts are among these techniques. Bond (2014) states that flowcharts assist developers "to operationalize their ideas" (p. 164). McCarthy et al. (2016) view flow chart as "highly detailed but complex" (p. 358). The advantage of using flowchart is listed by Ravi K. Walia (n.d.) as follows.

> - Flowchart is an excellent way of communicating the logic of a program.
> - Easy and efficient to analyse problem using flowchart.
> - During program development cycle, the flowchart plays the role of a blueprint, which makes program development process easier.
> - After successful development of a program, it needs continuous timely maintenance during its operation. The flowchart makes program or system maintenance easier.
> - It is easy to convert the flowchart into any programming language code. (Ravi K. Walia, n.d., p. 4)

In this research, the main steps will be shown using general flowcharts. Then sub-flowcharts will provide details of the main steps. The technique of using flowcharts to create DSR methods artefact will continue during the entire development of a prototype that will be presented in section 6.3.4. As a demonstration of the usage of flowcharts, this section will discuss and present the ways that flowcharts will store, retrieve and display examination questions. Furthermore, this section shows the implementation of the steps given in these flowcharts. The formation of flowcharts and implementation will focus on creating and displaying five questions. Figure 6.10 shows the creation of these questions. The explanations of the entities in this flowchart are shown in Table 6.7

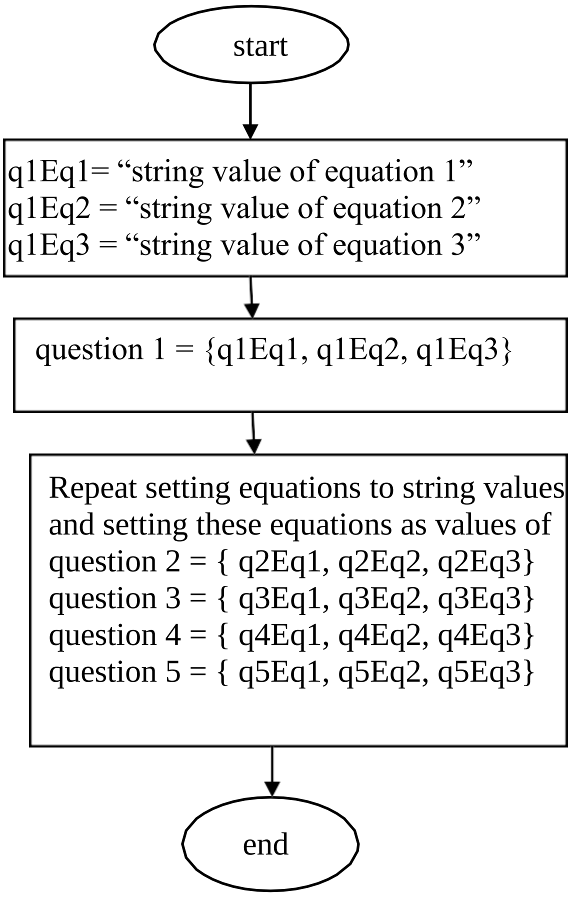


Figure 6.10 Questions setting

Table 6.7 entities in Figure 6.10 flowchart

Entities in the flowchart of question setting	Description
q1Eq1	A string variable that holds the first equation of a question
q1Eq2	A string variable that holds the second equation of a question
q1Eq3	A string variable that holds the third equation of a question
question 1	A string array variable that stores question 1 which contains linear system that has three equations.

The flowchart in Figure 6.10 is implemented by code that is shown in Table 6.8.

Table 6.8 Implementation of steps in Figure 6.10 flowchart

Java codes for steps in Figure 6.10 flowchart

```
public class QInArrayFormsGui
{

   StringBuffer q1Equation2 = new StringBuffer ("2x- 8y + 8z = -2");
   StringBuffer q1Equation3 = new StringBuffer ("-6x + 3y + 15z = 9");

   //Question 2 equations
    StringBuffer q2Equation1 = new StringBuffer ("2x  - y + z =1");
    StringBuffer q2Equation2 = new StringBuffer ("3x + 2y -4z = 4");
    StringBuffer q2Equation3 = new StringBuffer ("-6x + 3y -3z = 2");

   //Question 3 equations
    StringBuffer q3Equation1 = new StringBuffer ("x - y + 2z = -3");
    StringBuffer q3Equation2 = new StringBuffer ("4x + 4y - 2z = 1");
    StringBuffer q3Equation3 = new StringBuffer ("-2x + 2y - 4z = 6");


   //Question 4 equations
    StringBuffer q4Equation1 = new StringBuffer ("x + 2y + z = 5"); StringBuffer
    q4Equation2 = new StringBuffer ("3x + 2y + 4z = 17"); StringBuffer q4Equation3
    = new StringBuffer ("4x + 4y + 3z = 26");


     //Question 5 equations
    StringBuffer q5Equation1 = new StringBuffer ("2y + z = 5");
    StringBuffer q5Equation2 = new StringBuffer ("3x + 4z = 17");
    StringBuffer q5Equation3 = new StringBuffer ("4x + 4y  = 26");

          StringBuffer [] question1;
          StringBuffer [] question2;
          StringBuffer [] question3;
          StringBuffer [] question4;
          StringBuffer [] question5;


 public QInArrayFormsGui()
   {
      //setting equations to each question
           question1  =  new  StringBuffer  []  {q1Equation1,q1Equation2,q1Equation3};
           question2  =  new  StringBuffer  []  {q2Equation1,q2Equation2,q2Equation3};
           question3  =  new  StringBuffer  []  {q3Equation1,q3Equation2,q3Equation3};
           question4  =  new  StringBuffer  []  {q4Equation1,q4Equation2,q4Equation3};
           question5 = new StringBuffer [] {q5Equation1,q5Equation2,q5Equation3};
   }
}
```

Before displaying a requested question, it is required to check for its existence and retrieve it. The flowchart in Figure 6.11 shows steps that fulfil this purpose.

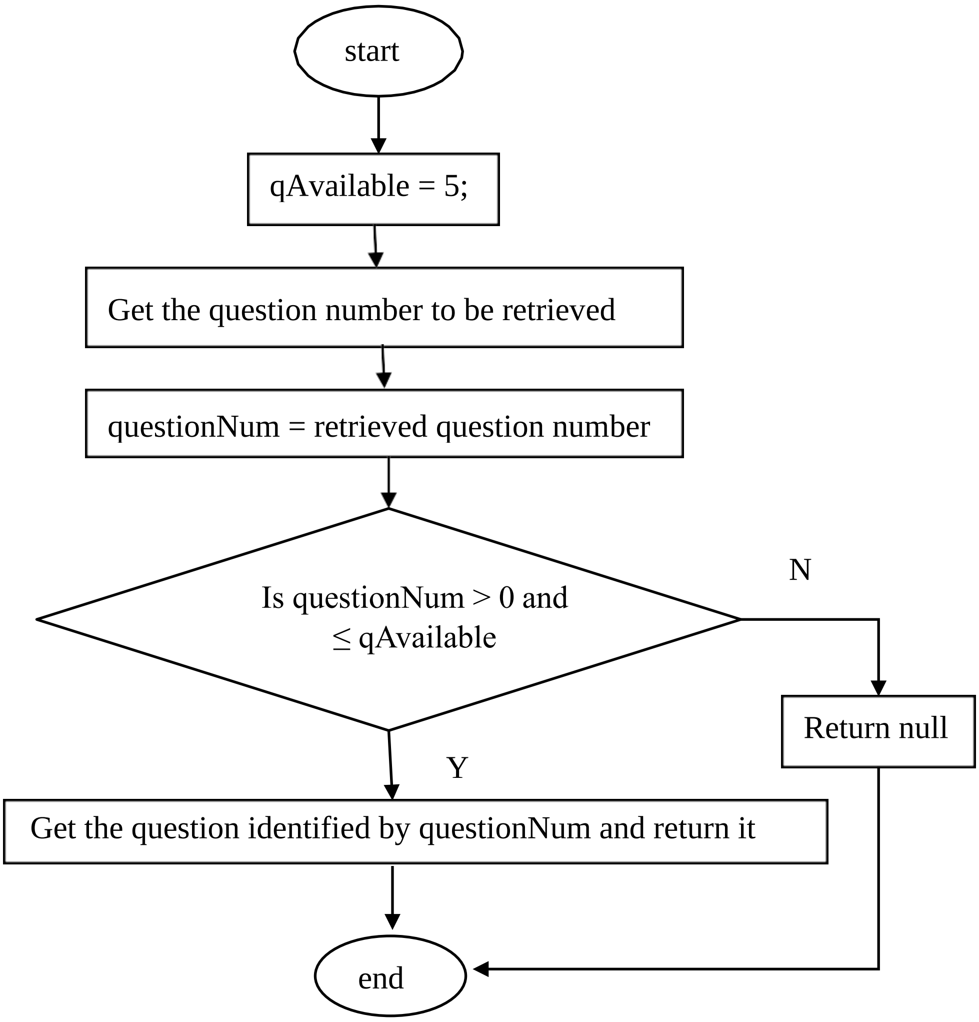


Figure 6.11 Retrieving a requested question

Table 6.9 shows the descriptions of the entities in Figure 6.11 flowchart.

Table 6.9 Entities in Figure 6.11 flowchart

Entities in the flowchart of question retrieving	Description
qAvailable	A variable that stores the value of five which is the number of questions that have been stored.
questionNum	The question number that identifies the question to be retrieved.
Is questionNum > 0 and ≤ qAvailable	The decision mechanism that checks for question number between 1 and 5. If the question number is between 1 and 5, get that question which is identified by the given number. Otherwise, return the value of 'null' to indicate that the question that has been requested is not available.

The code in Table 6.10 has not implemented the exact steps that have been shown in the Figure 6.11 flowchart. However, the code achieves the purpose of the flowchart which is retrieving the requested question. The ‘qAvailable’ constant value has not been used in this code. Instead the ‘questionNum’ value is compared with constant values 1 to 5 and the decision is made based on the outcome of the comparison.

Table 6.10 Implementing steps in Figure 6.11 flowchart

Java codes for steps in Figure 6.11 flowchart

```
public StringBuffer [] getQuestion(int questionNum)
  {
     if (questionNum == 1)
          return question1;
     if (questionNum == 2)
         return question2;
     if (questionNum == 3)
         return question3;
     if (questionNum == 4)
         return question4;
     if (questionNum == 5)
         return question5;
     else
         return null;
  }
```

The next step after retrieving the requested question is to prepare the question for display. The flowchart in Figure 6.12 starts from the point where the question has already been retrieved and saved in a string array variable called ‘question [ ]’.

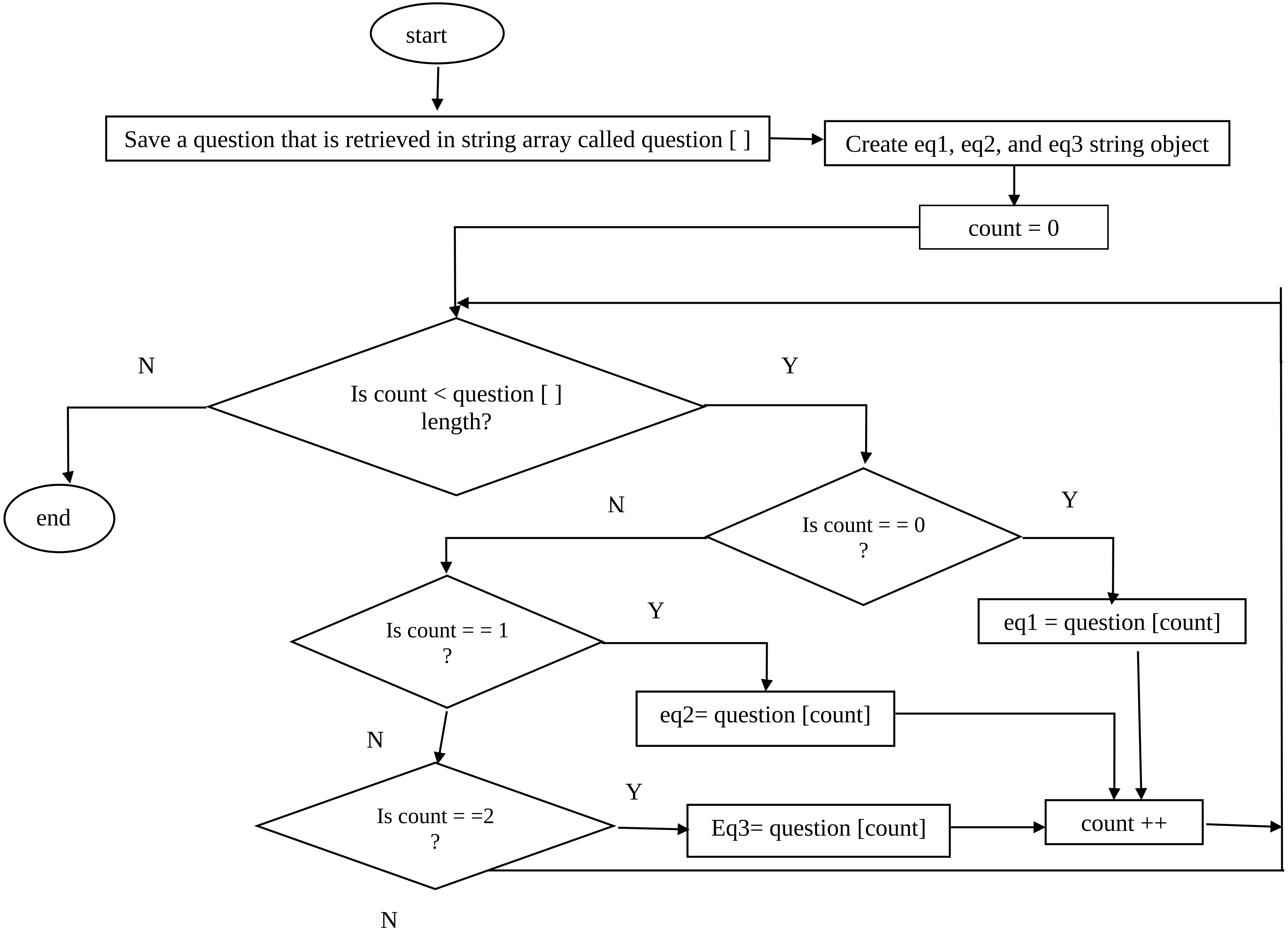


Figure 6.12 Preparing question for display

The descriptions of the entities that are shown in Figure 6.12 flowchart are given in

Entities in the flowchart of preparing the question for	Descriptions
Save a question that is retrieved in string array called question [ ]	This is a process that stores the question that has been retrieved using the steps in Figure 6.11. The question is stored in the string array variable called question [ ].
Create eq1, eq2 and eq3 string object	This is the process that creates three string variables that will be used to store the three equations that are in the retrieved question.
count = 0	This is the counter that is used to track a string location in the array that stores equations in the retrieved question. The counter is set to zero. In Java array the first position is indexed by zero.
Is count < question [ ] length?	The decision to be made is based on the length of the array which is, in this case, the number of equations in the question and the value in the count variable. After checking all locations in the array, the process ends after saving equations in the eq1, eq2 and eq3 string variables.

.

Table 6.11 Entities in Figure 6.12 flowchart

Entities in the flowchart of preparing the question for display	Descriptions
Save a question that is retrieved in string array called question [ ]	This is a process that stores the question that has been retrieved using the steps in Figure 6.11. The question is stored in the string array variable called question [ ].
Create eq1, eq2 and eq3 string object	This is the process that creates three string variables that will be used to store the three equations that are in the retrieved question.
count = 0	This is the counter that is used to track a string location in the array that stores equations in the retrieved question. The counter is set to zero. In Java array the first position is indexed by zero.
Is count < question [ ] length?	The decision to be made is based on the length of the array which is, in this case, the number of equations in the question and the value in the count variable. After checking all locations in the array, the process ends after saving equations in the eq1, eq2 and eq3 string variables.

The steps in Figure 6.12 flowchart are implemented by the code in

Table 6.12. This code extracts equations in the array and saves each equation as a string in eq1, eq2 and eq3 objects.

Java codes for implementing steps in Figure 6.12 flowchart

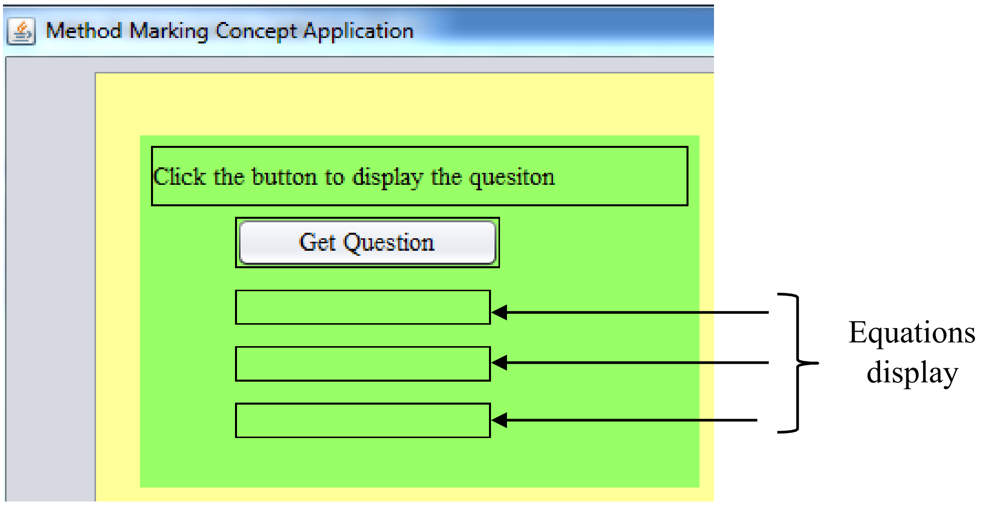
```
public void setEquations()
   {

      System.out.println("======= ");
      for ( int i = 0; i < question.length; i++)

      {

         if(i ==0)
         {
            //Get equation 1 from the question

            equation1 = question[0].toString();}

         if(i ==1)
         {
             //Get equation 2 from the question
            equation2 = question[1].toString();
        }

         if(i ==2)
         {
             //Get equation 3 from the question
            equation3 = question[2].toString();
         }

      }
}//End of setEquation() method
```

Table 6.12 Implementing steps in Figure 6.12 flowchart

Java codes for implementing steps in Figure 6.12 flowchart

```
public void setEquations()
   {

      System.out.println("======= ");
      for ( int i = 0; i < question.length; i++)

      {

         if(i ==0)
         {
            //Get equation 1 from the question

            equation1 = question[0].toString();}

         if(i ==1)
         {
             //Get equation 2 from the question
            equation2 = question[1].toString();
        }

         if(i ==2)
         {
             //Get equation 3 from the question
            equation3 = question[2].toString();
         }

      }
 }//End of setEquation() method
```

To display a question, consider the GUI templet that is shown in Figure 6.13. This action starts when the 'Get Question' button is clicked. The code that is executed when this button is clicked is shown in Table 6.13.



Click the button to display the quesiton

Figure 6.13 Question display request initiator

In this code, the ‘psTask.displayQuestion (this)’ statement starts the process of requesting a question display. The flowchart in Figure 6.14 shows the steps that are implemented in this method.

Table 6.13 Codes to display a question

Java code of ‘Get Question’ button shown in Figure 6.13

```
private void psQuestionRequestButtonActionPerformed(java.awt.event.ActionEvent evt)
 {

        //
        psTask = new PsTasks();
        //Reads the question from the object that store it
        //and assign the equations in the question to JLabel objects
        //the this keyword passed to the displayQuestion method is the
        //Frame that contains the button that this ButtonActionPerformed belongs to
        psTask.displayQuestion(this);

        }
```

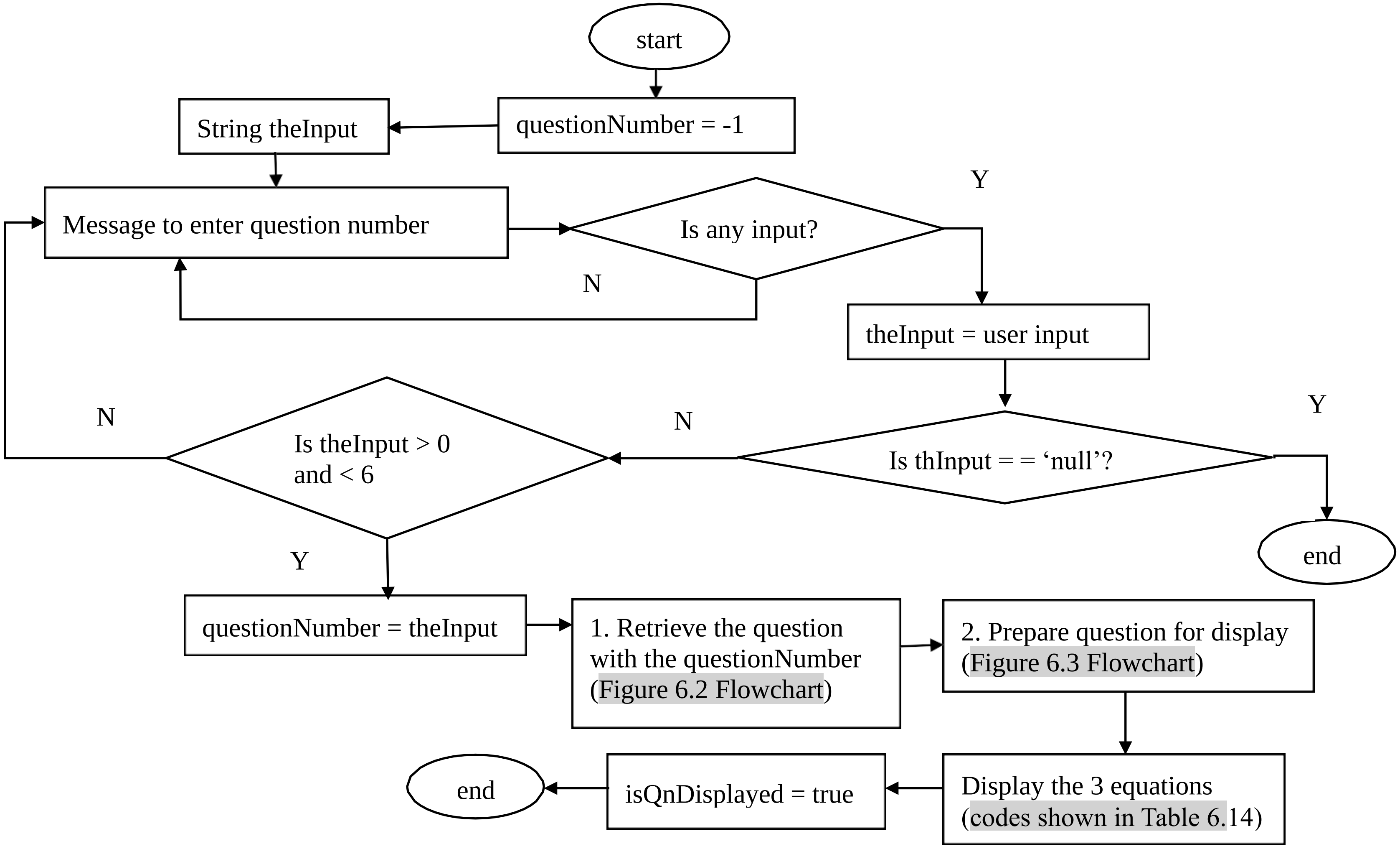


Figure 6.14 Question retrieve, prepare and display process

Table 6.14 contains the descriptions of the entities in the flowchart that has been displayed in Figure 6.14.

Table 6.14 Entities in Figure 6.14 flowchart

Entities in the flowchart of question retrieving, preparing and displaying	Description
questionNumber = -1	A variable that stores the question number of the requested question. Initially it is set to -1 to show that this value as the question number is invalid.
String theInput	A variable that stores the value that is entered by a user.
Message to enter question number	A message for the user
Is any input?	A decision-making means that checks if the user has responded to the message to enter a question number. If no input from the user, keep displaying the message that prompts the user to input a question number. If any input is entered, save this value in the 'theInput' variable. If the theInput value is 'null' the entire steps presented by the flowchart ends. The 'null' value indicates the cancelling of the request for a question display. Otherwise the other action is to check for the value in the theInput variable for further decision to be made.
Is theInput > 0 and < 6	This decision making means checks for the value of theInput variable to verify if it is between 1 and 5. If the number is not in the range, continue displaying the message to prompt the user to enter the correct number or quit the process of requesting a question.
questionNumber = theInput	If the value in the theInput variable is between 1 and 5, save this value in the questionNumber variable.
Retrieve the question with the questionNumber (Figure 6.11 Flowchart)	This process is performed by steps in the flowchart of Figure 6.11.
2 .Prepare question for display (Figure 6.12 Flowchart)	This process is implemented in the steps that have been shown in Figure 6.12 flowchart.
Display the 3 equations (codes shown in Table 6.15)	This will be shown as code (highlighted grey) in Table 6.15.
isQnDisplayed = true	A flag to indicate that the requested question has been displayed. Therefore, set this flag to true. This flag was initially set to false before displaying the requested question.

The implementation of steps in the flowchart that is shown in Figure 6.14 is shown in the code that is displayed in Table 6.15.

Table 6.15 Code for implementation of the flowchart in Figure 6.14

Java code for steps in Figure 6.14 flowchart

```
  public void displayQuestion(WorkingsProgressFrame obj)
   {

     int questionNumber =-1;//to hold the integer value of the question number to be
displayed.
      String theInput; //to store the input which is the question number
     /*
      the while (true) loop continue displaying the dialog box until
      a valid input is entered or the cancel button is pressed.
      */
      while(true)
        {

          //Display the dialog box and read the input
            theInput = JOptionPane.showInputDialog(null, "Please enter 1-5.");
          if(theInput == null)//to check if the cancel button pressed
          {
             System.out.println("Test: the cancel button pressed");
           /*
             the cancel button stops executing the rest of the codes
             in the method and return to where the  method is called from
             */
            return;
          }
          else
          {
             if (theInput.matches("\\d"))
                   {
                        questionNumber = Integer.parseInt(theInput);
                        qNum = Integer.parseInt(theInput);
                        if ( questionNumber >0 &&  questionNumber <6)
                        {
                           System.out.println("Test: the question number is: " +
questionNumber);
                           break;
                        }
                   }
          }//end of else block

                }//end of while loop
```

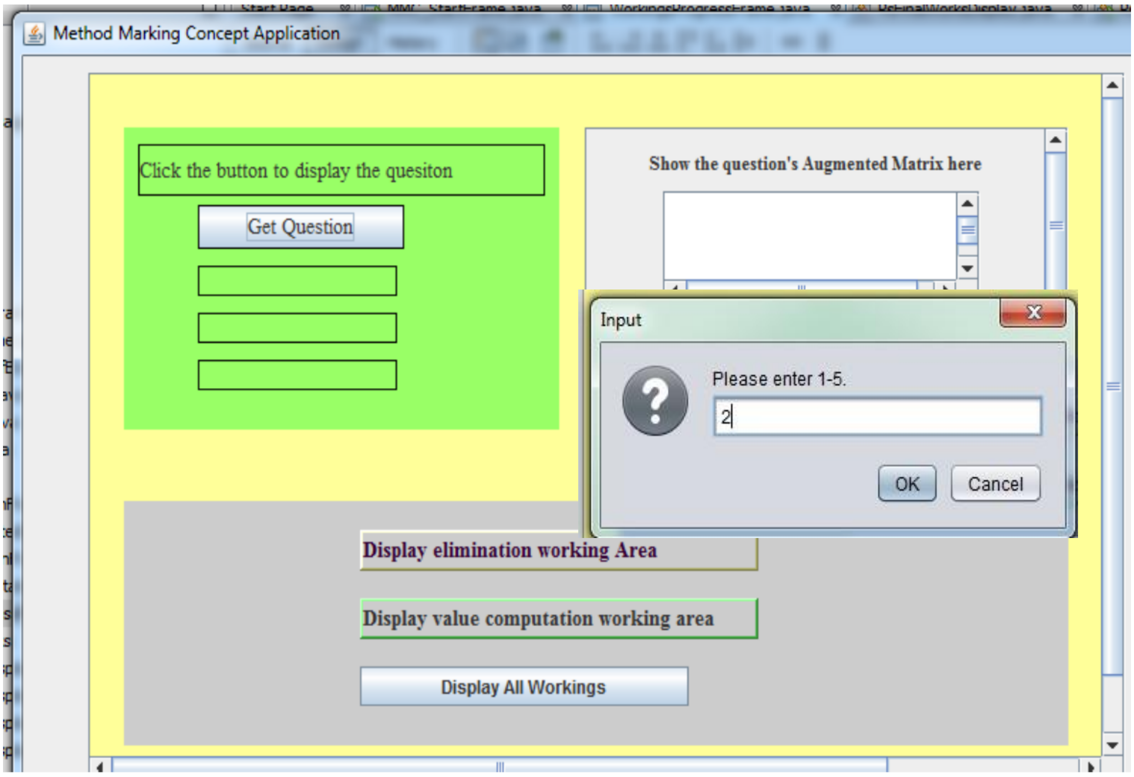
Java code for steps in Figure 6.14 flowchart

```
      matForm = new MatrixFormationGui (questionNumber);
      matForm.setEquations(); //separate the requested question into equations
 //Retrieve equations in the question and assign/display them as text in label objects

       obj.getEquation1Lbl().setText(matForm.getEquation1());
       obj.getEquation2Lbl().setText(matForm.getEquation2());
       obj.getEquation3Lbl().setText(matForm.getEquation3());

 //   make sure that question has been displayed and remained displayed
 //isQnDisplayed flag reports this if true during the period of the question solving
 //period.
        isQnDisplayed = true;

  }
```

The execution of the code in Table 6.15 performs the following actions.

1. Display a dialog box.
2. Prompt a user to enter a question number to be displayed.
3. Search for the requested question and display it.

The first and second steps are shown in Figure 6.15 while the third step is shown in

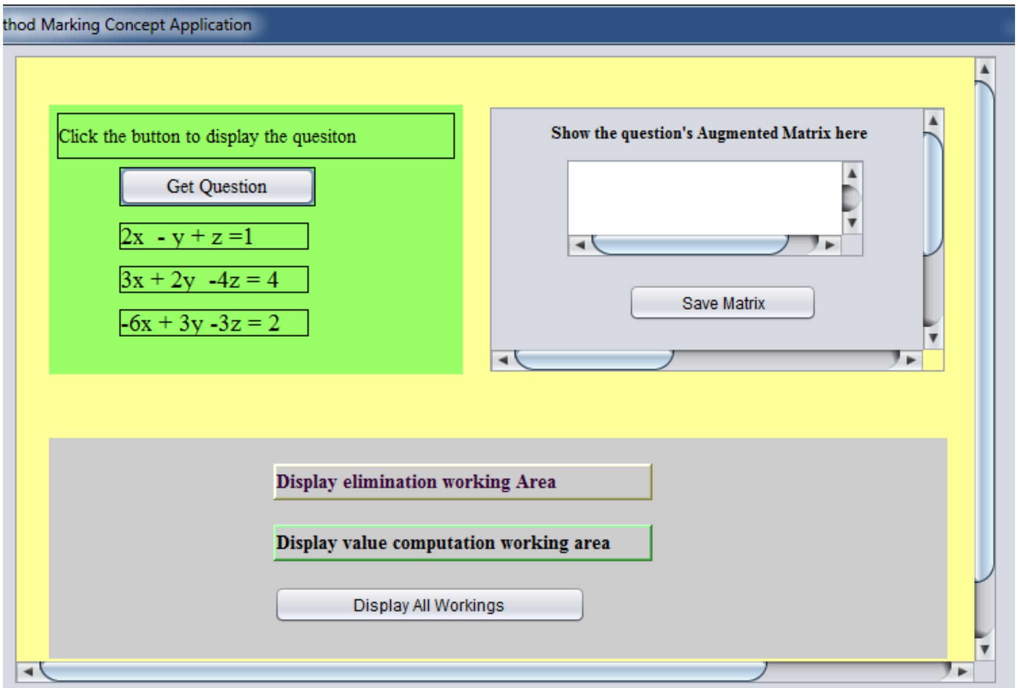


Please enter 1-5.

2

Figure 6.15 Requesting question 2 display

Figure 6.16.



2x - y + z =1

3x + 2y -4z = 4

-6x + 3y -3z = 2

Figure 6.16 Question equations display

6.3.4 Method marking concept (MMC) prototype.

An ‘instantiation’ is a useable product that might be abstract or physical and it is produced from given designs (Venable, 2006). An instantiation “operationalize constructs, models, and methods” (March & Smith, 1995, p. 258). According to March and Smith, it is not necessary to produce fully operational or usable constructs, models and methods to develop an instantiation. This is due to “an instantiation may actually precede the complete articulation of its underlying constructs, models, and methods” (March & Smith, 1995, p. 258). In this research, however, both approaches have been considered. After understanding the research problem from the findings of the analyses of scripts and the literature review, some parts of the prototype have been created. However, the other parts of the prototype were created after the creation of constructs, models and methods. In this research, when presenting these artefacts, the order that has been followed is: constructs, models, methods and prototype (instantiation).

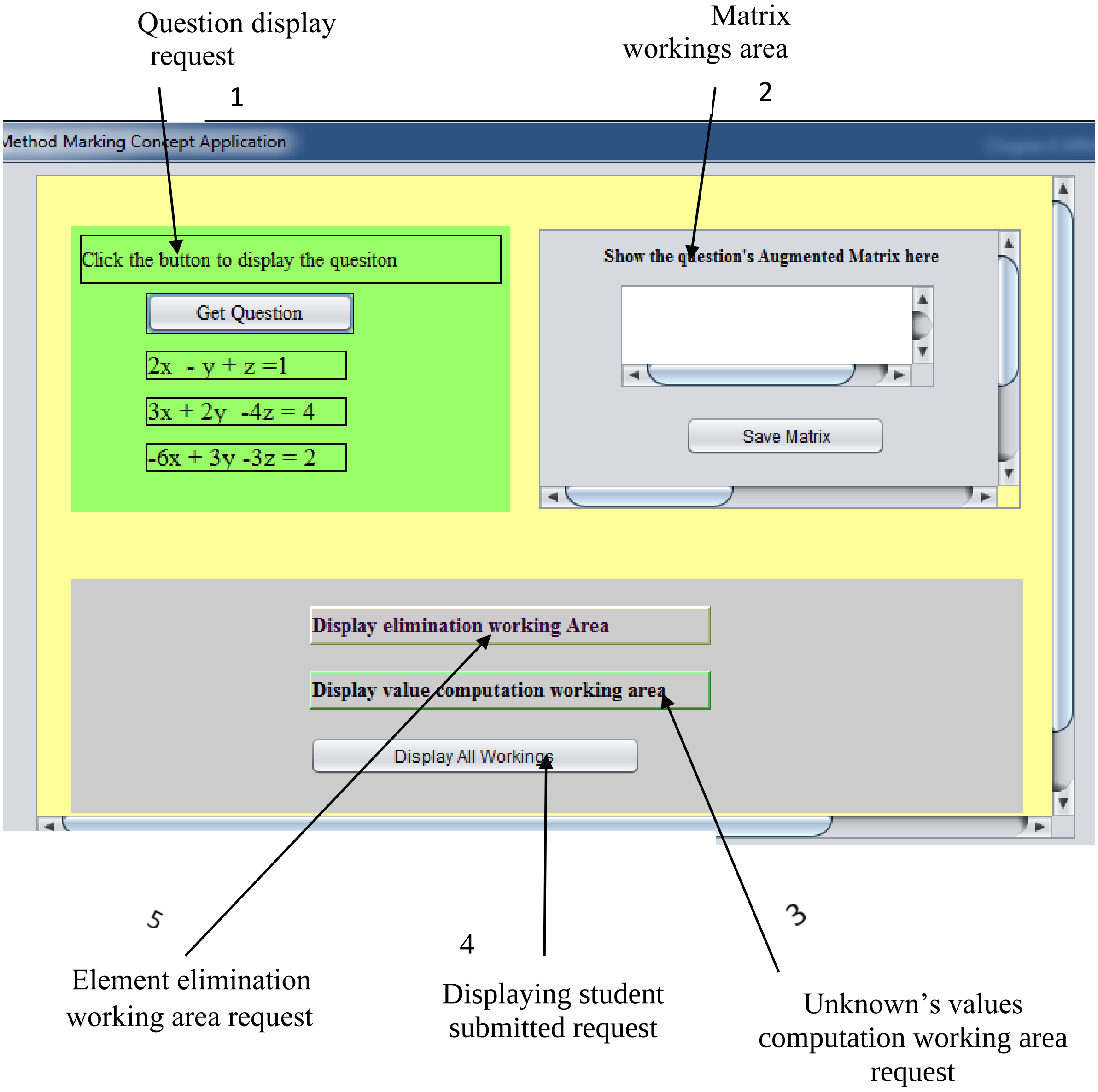
The models that have been shown in section 6.3.2 consist of two parts which are concerned with students and assessors' tasks. However, this prototype focuses on tasks which allow students to perform the following actions.

1. Request and display an examination question.
2. Record their workings and submit.
3. Display their workings for modification.

The first task of creating this prototype has started during the demonstration DSR methods artefacts in section 6.3.3. This means the development of the prototype has been in progress from the previous section. This section continues the explanations and development of this prototype. Furthermore, the evaluation of the given code and results of that evaluation are discussed at the end of the tasks that are performed through interaction of students and this prototype. The strategy that will be used here will be explaining tasks that are performed by each part in this prototype. To employ this strategy, user interfaces (UI) that consist of windows and their components will be displayed and functions of each component will be explained. These explanations will also include interactions of users, input required and resulting output. However, only selected visual components (windows and/or dialog boxes) will be used when giving these explanations.

6.3.4.1 MMC prototype home user interface.

The first window for students' interaction is the 'home window'. This window is shown in Figure 6.17 with five arrows pointing to five different components. The component that is pointed to by the arrow 1 performs the tasks of retrieving and displaying a question. The details of steps (DSR methods) required to produce this component have been presented in section 6.3.3. The second component that is pointed to by the arrow 2 is the working area that allows students to enter a matrix and submit it for assessment. The third component that is pointed to by the arrow 3 allows students to request a window that permits students to perform the task of computing values of ‘unknowns’. The component that is pointed to by the arrow 4 allows students to display the already submitted workings for the purpose of updating or deleting.



2x - y + z =1

3x + 2y -4z = 4

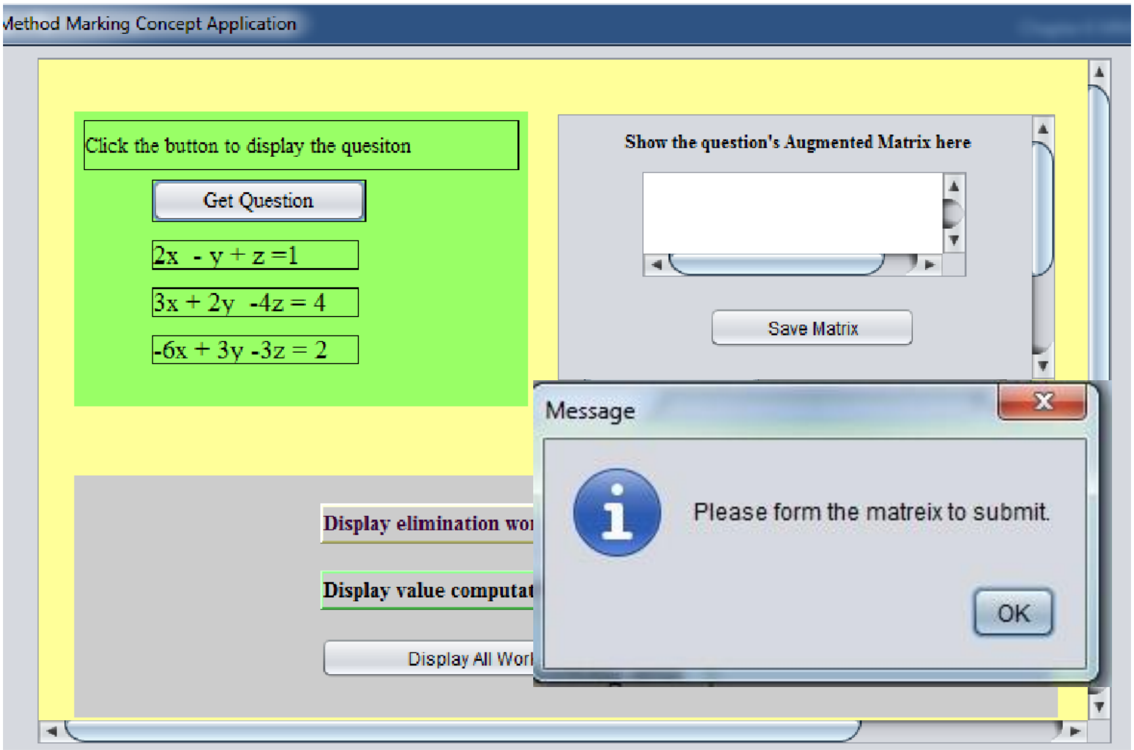
-6x + 3y -3z = 2

Figure 6.17 MMC prototype home window

The last component that is pointed to by the arrow 5 allows students to display a window that is used to enter workings of eliminating elements. The details of all tasks that are performed using the utilities that have been shown in this home window will be given in the following sections. However, the details of the task of displaying a question will not be covered because this issue has already been presented in section 6.3.3.

6.3.4.2 Matrix formation process.

The actions that students perform when producing a matrix have been shown in the model that has been displayed in Figure 6.4 and explained in section 6.3.2.2. For this operation to be successful, students are required to display a question using the ‘Get Question’ button. Clicking on ‘Save Matrix’ button will display an error and request for an input if either a question is not displayed, or a matrix has not been entered in the text area. Figure 6.18 shows this scenario. However, after a question is displayed, all types of inputs other than empty working are accepted as a legitimate input that is considered as a matrix. Figure 6.19 shows the displayed question and matrix.



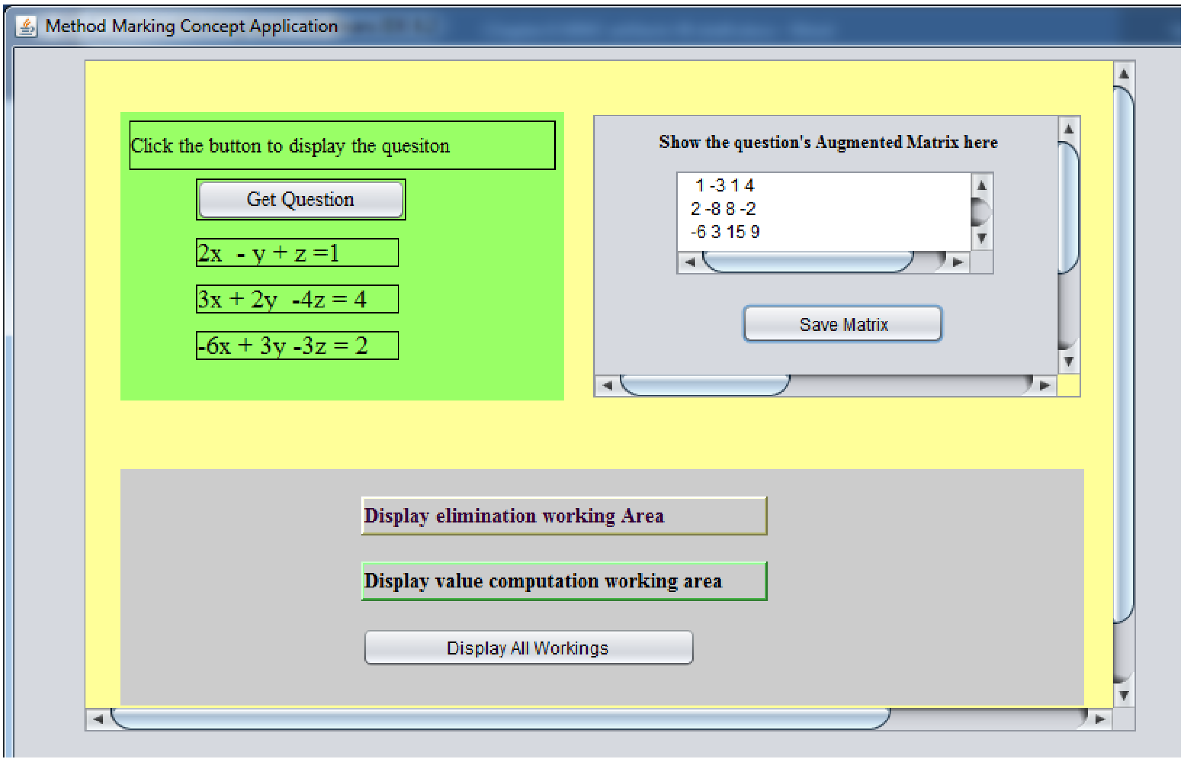
2x - y + z =1

3x + 2y -4z = 4

-6x + 3y -3z = 2

Please form the matreix to submit.

Figure 6.18 Empty matrix workings submission alert message



2x - y + z =1
3x + 2y -4z = 4
-6x + 3y -3z = 2

1 -3 1 4
2 -8 8 -2
-6 3 15 9

Figure 6.19 Question and matrix in workings area

The submitted working is saved in a file. To perform this task, the prototype (1) gets the question number that identifies the matrix to be saved and (2) checks for the existence of a file that will store the working. (3) If a file is not available, the prototype creates a new file, and (4) writes this file's name on the first line of the file, then goes to a new line and writes the submitted matrix. To check the correct execution of steps (tasks) during the formation of a matrix, different methods have been used to evaluate this prototype. One method is to physically open the file that stores the student workings and check its content as well as making observations of student interactions and the prototype's responses. The other method is shown in Table 6.16.

This table shows some of important events or actions that have been captured during the execution of code that has been used to form a matrix. The content of the table has been produced by 'System.out.println' Java language statement. This statement has been embedded in different relevant parts of the code. This statement has been extensively used as a form of testing/evaluating code to check whether it fulfils the purpose of its creation. The statement prints messages that are passed to it. The explanations of these message are given in the right column of Table 6.16.

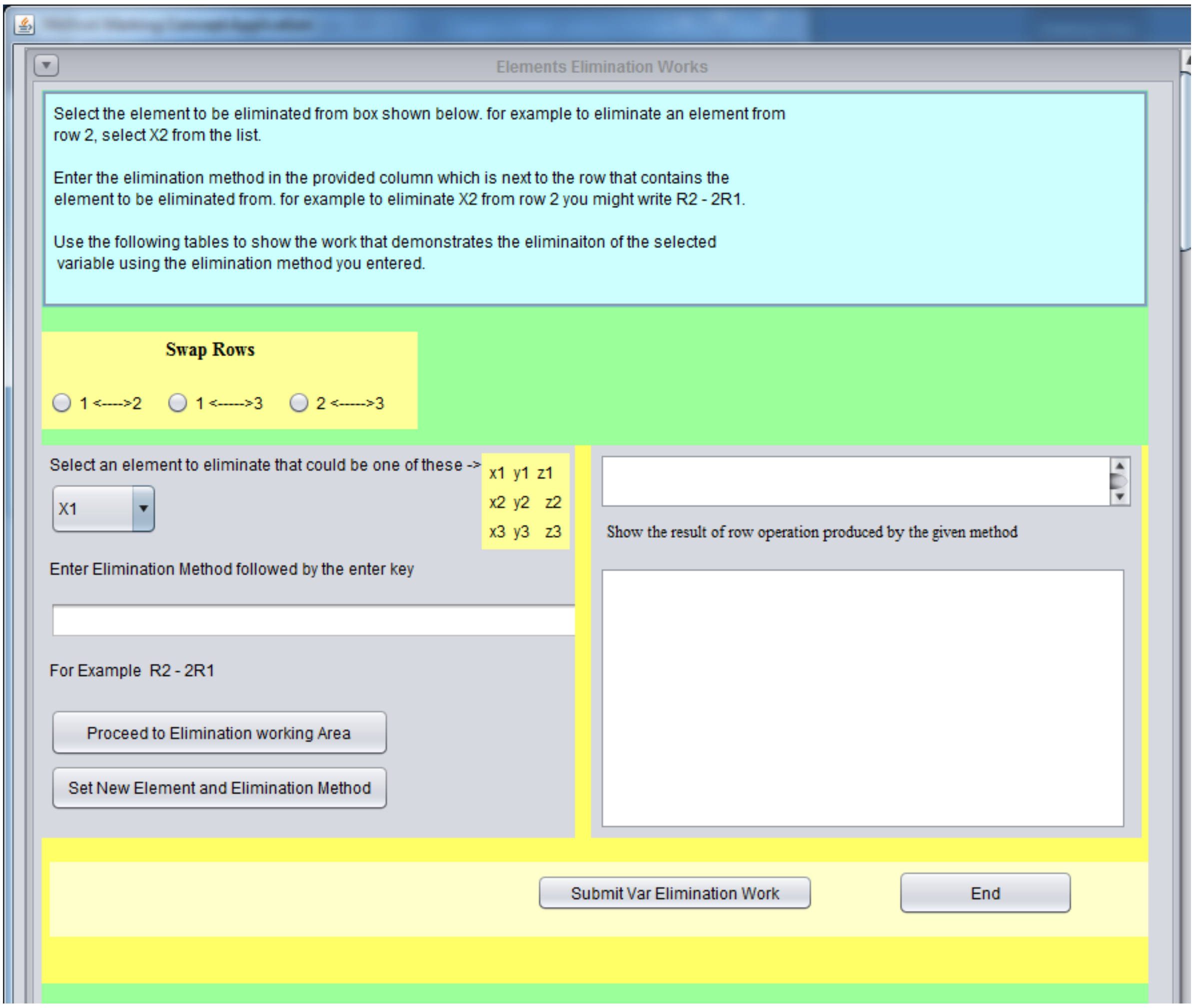
Table 6.16 Matrix formation evaluation result

Messages printed by the embedded 'System.out.println' statement during matrix formation process	Explanations
Test: the question number is: 1	This is the question number which a student has requested to produce a matrix.
Test: The question is displayed	This message indicates that the question has been retrieved and displayed.
Test: The working area is not empty.	This message shows that the student has entered a value in the matrix workings area.
Test: the option is yes	This is the message that shows that the student has submitted his/her working.
the file path: .\src\problemSolversWorksFiles\q1\matrixFormation.txt	The message that shows the file path where the submitted working has been stored or going to be stored.
file named .\src\problemSolversWorksFiles\q1\matrixFormation.txt already exists	This is the message that shows the file which is named 'matrixFormation.txt' has already been created.
The file content:	The message to show that the following information is the content of the file.
matrixFormation.txt 1 -3 1 4 2 -8 8 -2 -6 3 15 9	The file content starts with the file name which is followed by a new line. After new line the value of the matrix in the file is displayed.

The data in this table and the visual presentation of the matrix formation process show that this part of the porotype correctly performs a matrix formation task. The next step is the element elimination process that is presented in section 6.3.4.3.

6.3.4.3 Element elimination process.

The model in Figure 6.20 shows actions that students take to record their workings. One of these workings is eliminating elements from a matrix. This part of the prototype allows students to record this working. In Figure 6.17 the arrow labelled with 5 points to the component with a ‘Display elimination working area’ title. This component displays a window that is displayed in Figure 6.20.



Select the element to be eliminated from box shown below. for example to eliminate an element from row 2, select X2 from the list.

Enter the elimination method in the provided column which is next to the row that contains the element to be eliminated from. for example to eliminate X2 from row 2 you might write R2 - 2R1.

Use the following tables to show the work that demonstrates the eliminaiton of the selected variable using the elimination method you entered.

Select an element to eliminate that could be one of these ->

X1

Enter Elimination Method followed by the enter key

For Example R2 - 2R1

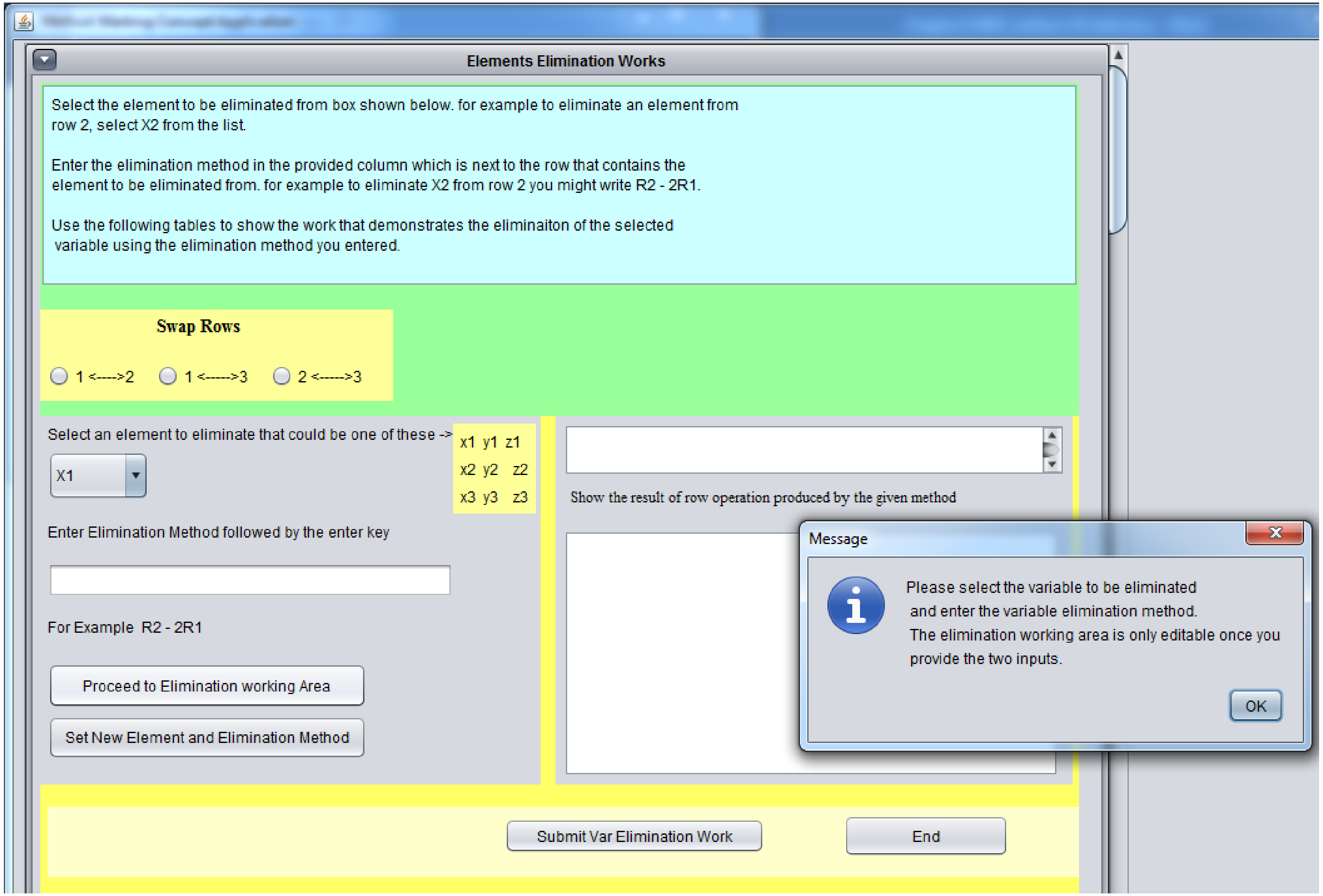
Show the result of row operation produced by the given method

Figure 6.20 Elements elimination workings area and tools

This window contains the components that are shown below.

- An area that is shown at the top of the window displaying an instruction.
- ‘Radio buttons’ that are designed to monitor if any occurrence of swapping equations happens. This functionality has not been implemented in this prototype.
- A ‘drop-down list’ to allow students to select an element to be eliminated.
- A ‘text field’ to enter an ‘elimination method’ for the selected element.

- A button with the ‘proceed to elimination working area’ title that triggers the following actions:
    - Checks if the student has selected an element to be eliminated and entered an elimination method.
    - If the two inputs are provided, displays them in the upper smaller text area.
    - Puts the focus of the cursor in the ‘working area’ which is the bigger text area.
    - If the two inputs have not been provided displays a message that prompts students to provide them. This scenario is shown in Figure 6.21.
- A ‘text area’ to display the selected element and elimination method.
- A text area for student to show their workings.
- A button with the ‘Set New Variable and Elimination Method’ title. This button clears the selected element and elimination method input components.
- A button with the ‘Submit Var Elimination Work’ title. If students try to submit empty workings, the prototype responds with the message that prompts students to enter their workings. The same as in matrix formation, the prototype accepts any input excluding empty or null value. This scenario is shown in Figure 6.22.



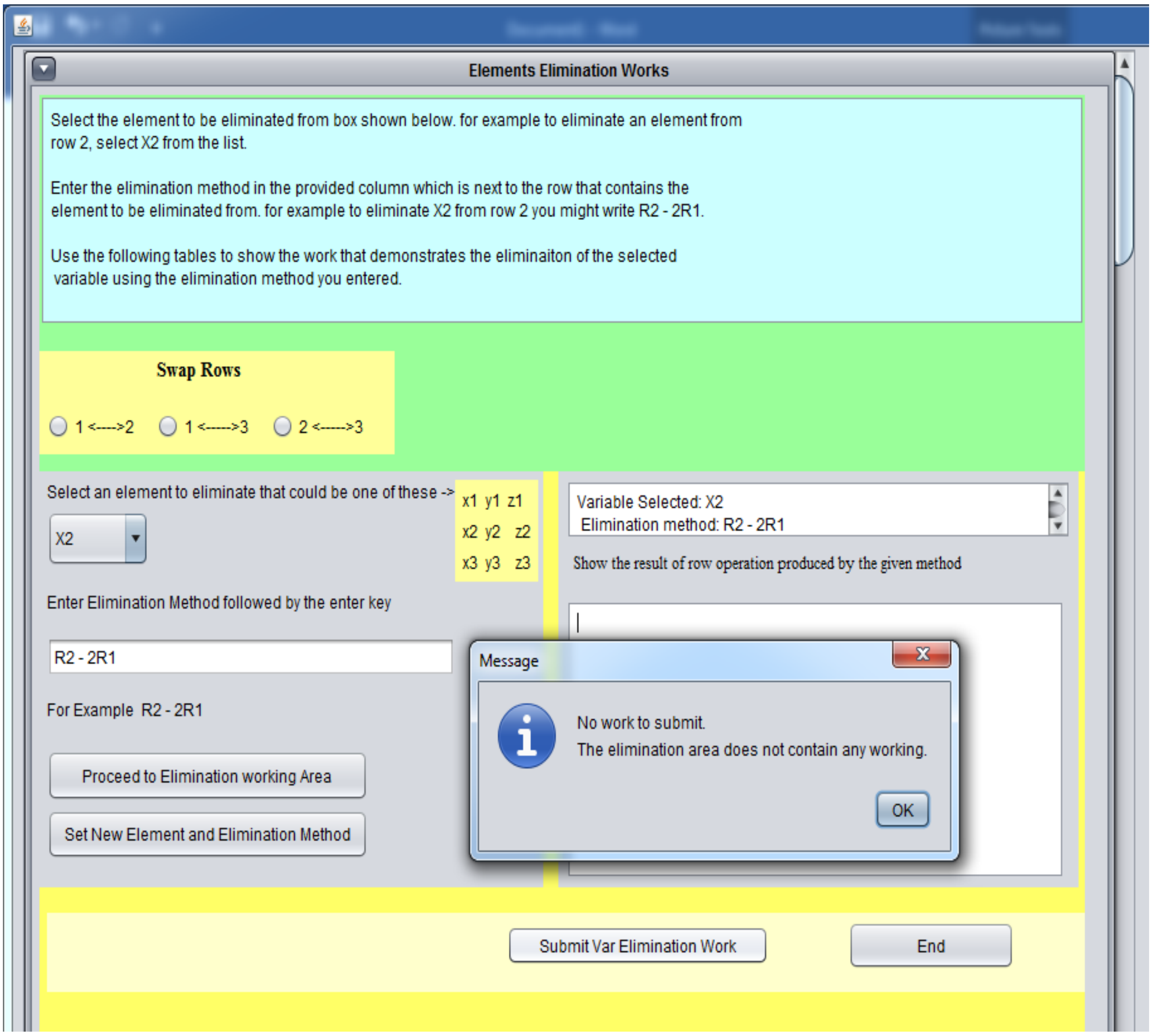
Select the element to be eliminated from box shown below. for example to eliminate an element from row 2, select X2 from the list.

Enter the elimination method in the provided column which is next to the row that contains the element to be eliminated from. for example to eliminate X2 from row 2 you might write R2 - 2R1.

Use the following tables to show the work that demonstrates the eliminaiton of the selected variable using the elimination method you entered.

Please select the variable to be eliminated and enter the variable elimination method. The elimination working area is only editable once you provide the two inputs.

Figure 6.21 Prompt to select element to be eliminate and to input elimination method



Select the element to be eliminated from box shown below. for example to eliminate an element from row 2, select X2 from the list.

Enter the elimination method in the provided column which is next to the row that contains the element to be eliminated from. for example to eliminate X2 from row 2 you might write R2 - 2R1.

Use the following tables to show the work that demonstrates the eliminaiton of the selected variable using the elimination method you entered.

Select an element to eliminate that could be one of these ->

X2

Enter Elimination Method followed by the enter key

R2 - 2R1

For Example R2 - 2R1

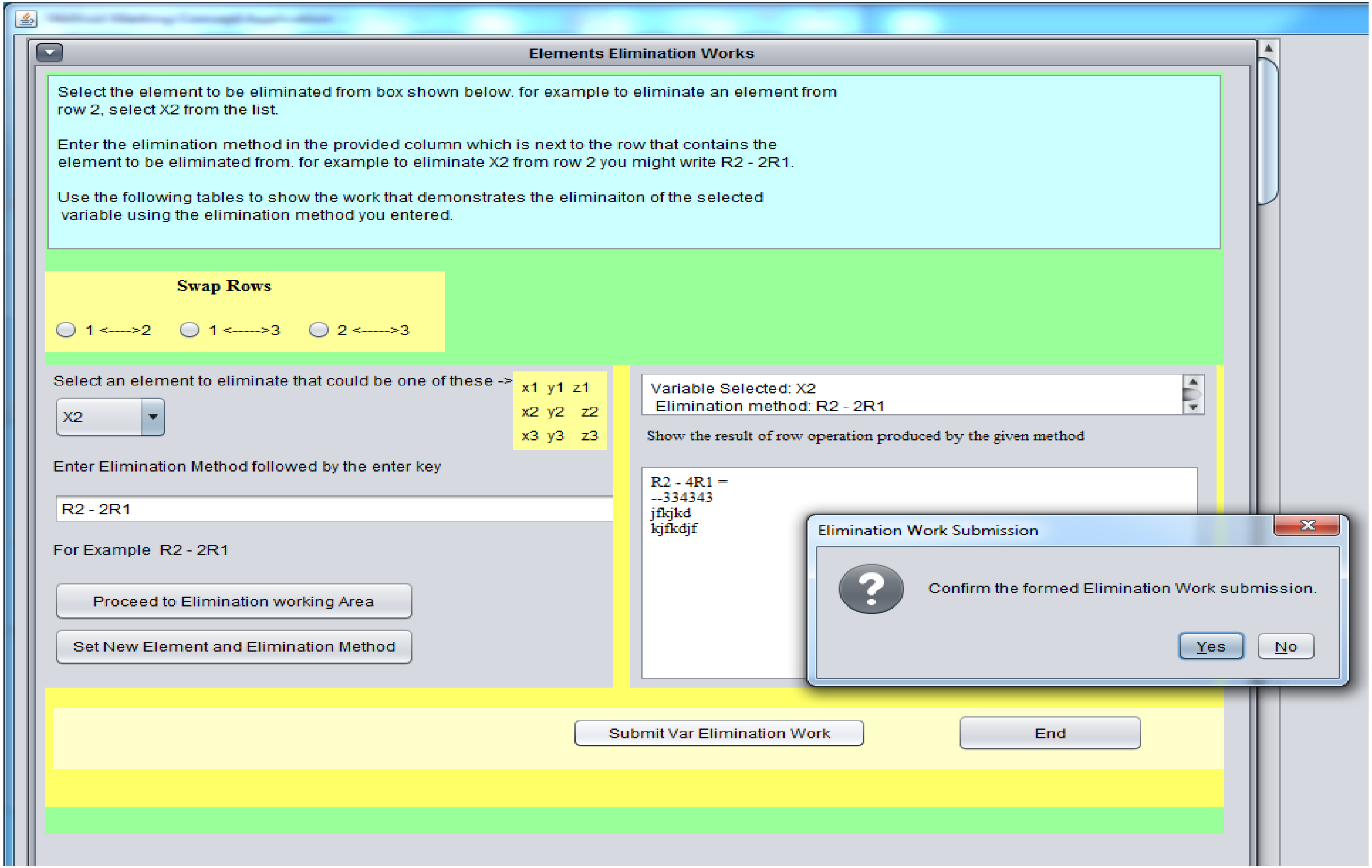
Variable Selected: X2
Elimination method: R2 - 2R1

Show the result of row operation produced by the given method

No work to submit.
The elimination area does not contain any working.

Figure 6.22 *Prompt to enter workings in working area*

In Figure 6.23 clicking of ‘submit Var Elimination Work’ button displays the dialog box that prompts students to confirm storing their workings. In Figure 6.24, the prototype prompts the student to eliminate more elements or quit the elimination process.



Elements Elimination Works

Select the element to be eliminated from box shown below. for example to eliminate an element from row 2, select X2 from the list.

Enter the elimination method in the provided column which is next to the row that contains the element to be eliminated from. for example to eliminate X2 from row 2 you might write R2 - 2R1.

Use the following tables to show the work that demonstrates the eliminaiton of the selected variable using the elimination method you entered.

Select an element to eliminate that could be one of these ->

X2

Enter Elimination Method followed by the enter key

R2 - 2R1

For Example R2 - 2R1

Variable Selected: X2
Elimination method: R2 - 2R1

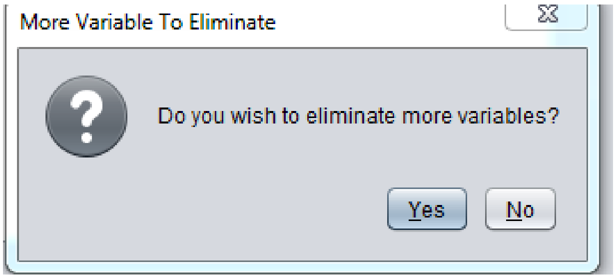
Show the result of row operation produced by the given method

R2 - 4R1 =
--334343
jfkjkd
kjfkdjf

Elimination Work Submission

Confirm the formed Elimination Work submission.

Figure 6.23 Elimination workings submitting confirmation


More Variable To Eliminate

Do you wish to eliminate more variables?

Figure 6.24 Prompt for more elimination of an element

The element elimination part of this prototype has successfully interacted with the student and provided the utility of allowing the student to perform the task of eliminating an element. The evaluation methods that have been used for this section are similar to the evaluation methods that have been used during the formation of matrix. Table 6.17 contains messages that have been produced at different stages of an element elimination.

Table 6.17 Element elimination evaluation result

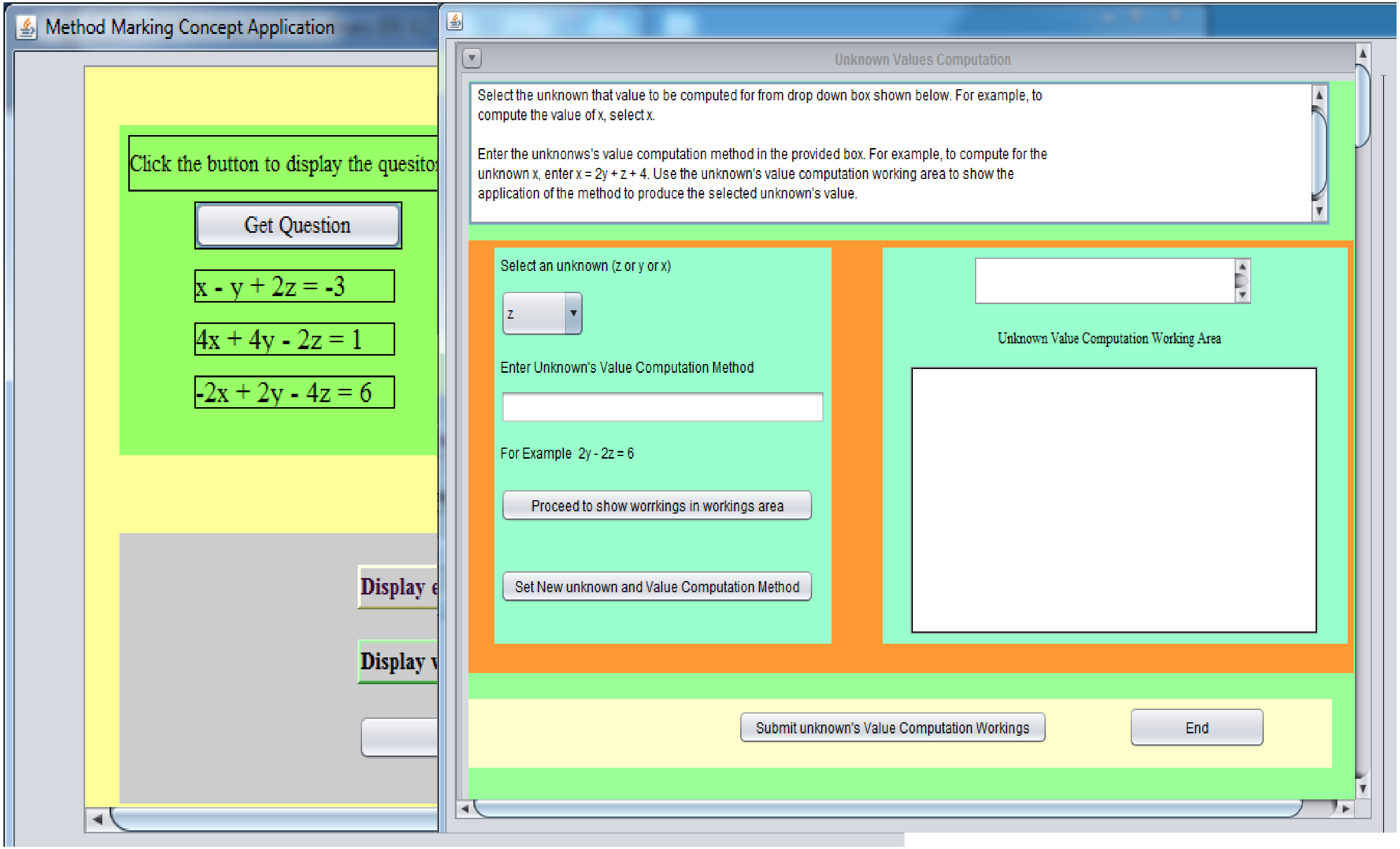
Messages printed by embedded 'System.out.println' 'Java language' statement during an element elimination process	Explanations
run: Test: the question number is: 3 ======= Test: The question displayed using psTask.getQNum method is 3	Stage 1 Before starting the process of eliminating an element, the question that contain the elements must be displayed first. The message informs that the question that has been displayed is question 3. This coincides with the number that is entered by the student. The 'psTask.getQNum' is the method that is used to retrieve the requested number.
Test: the selected element to be eliminated is X 2 and the provided elimination method is R2 - 4R1.	Stage 2 This is the first stage of showing an element elimination work. At this stage the element to be eliminated and the

Messages printed by embedded ‘System.out.println’ ‘Java language’ statement during an element elimination process	Explanations
	elimination method are provided by the student.
Test: workings contain: R2 - 4R1 = --334343 jfkjkd kjfkdjf Test: The elimination work is going to be written in the file	Stage 3 At this stage the prototype shows the details of the element elimination working that will be saved in a file.
Test: workings contain: R2 - 4R1 = --334343 jfkjkd kjfkdjf Test: The elimination work is going to be written in the file. Test: the option is yes the file path: .\src\problemSolversWorksFiles\q3\varX2ElimWork.txt file named .\src\problemSolversWorksFiles\q3\varX2ElimWork.txt already exists Test: Your work is successfully written in the file.	Stage 4 The final stage showing what is saved in the file. The working includes the elimination method and the details of using this method. The file path gives the following details: q3 is the folder that contains the question 3 answers. varX2ElimWork.txt is the file name that stores X2 element elimination work. The “already exists “message indicates the file has been already created. The last message reports the status of writing the content in the created or existing file.
Test: The file name is added to the file list. The file content: varX2ElimWork.txt Variable name = X2 Elimination Method = R2 - 4R1 R2 - 4R1 = --334343 jfkjkd kjfkdjf	The message in the left column is the content that is retrieved from the file that is used to store the elimination workings. Stage 5 The message “the file name is added to the file list” shows the files that are used to save student workings are saved in a file list. The file content starts with its name which is written on

Messages printed by embedded 'System.out.println' 'Java language' statement during an element elimination process	Explanations
	the first line of the file. The next line displays the name of the eliminated element. The next line shows the elimination method followed by the working which is the details of the usage of the elimination method.

6.3.4.4. Unknowns values computation process.

This part of this prototype allows students to show working that involves computing values of unknowns. The model in Figure 6.8 also covers this type of actions that is performed by students. In the 'home window' that has been displayed in Figure 6.17, the 'Display value computation working area' button displays the window that is displayed in Figure 6.25.



Click the button to display the quesito

x - y + 2z = -3

4x + 4y - 2z = 1

-2x + 2y - 4z = 6

Select the unknown that value to be computed for from drop down box shown below. For example, to compute the value of x, select x.

Enter the unknonws's value computation method in the provided box. For example, to compute for the unknown x, enter x = 2y + z + 4. Use the unknown's value computation working area to show the application of the method to produce the selected unknown's value.

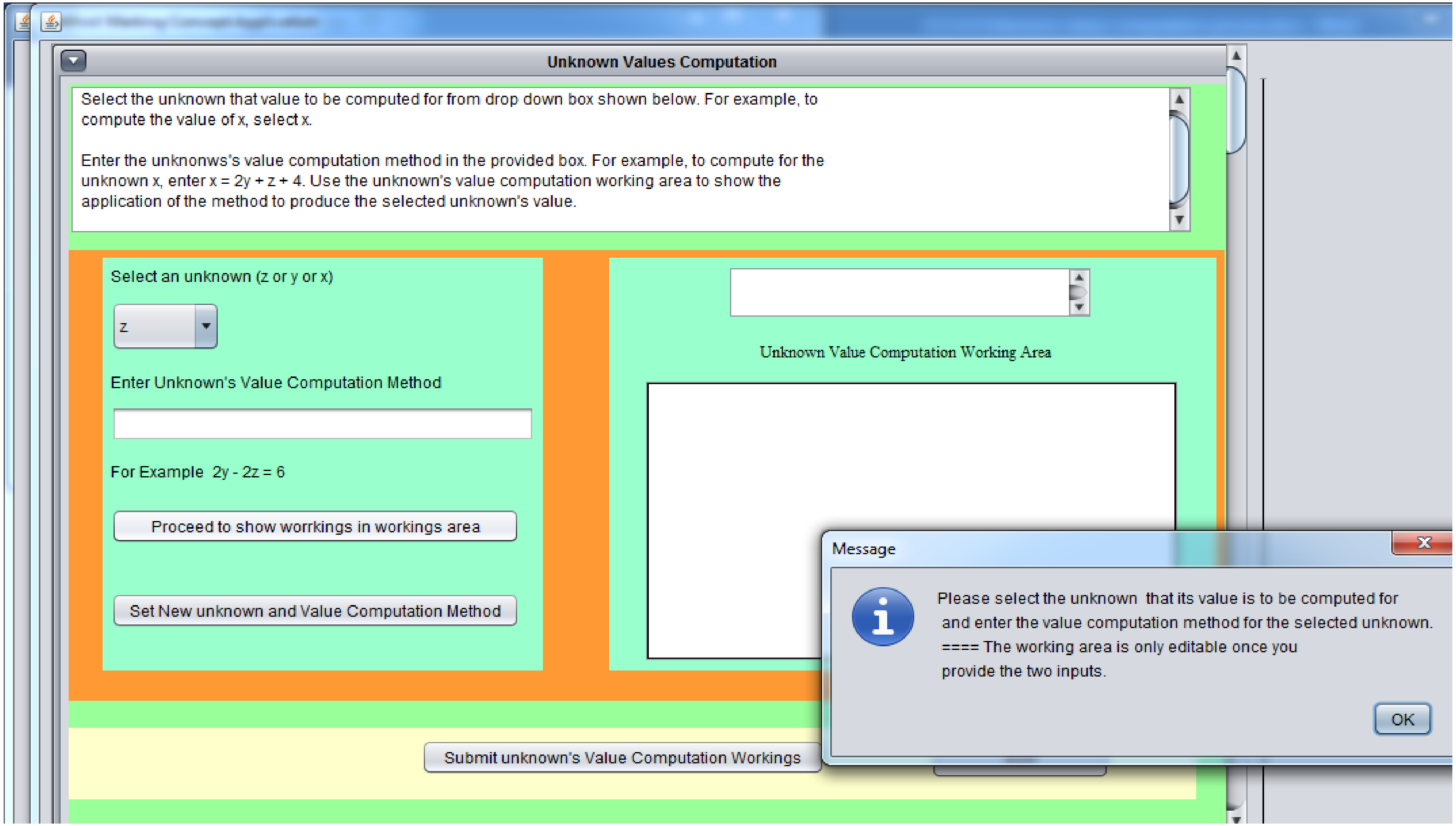
For Example 2y - 2z = 6

Figure 6.25 Unknown values computation process

The components in this window are listed below:

- A text area that displays the instruction that is shown at the top of the window.
- A drop-down list containing unknowns to be selected.
- A text field to read the value computation method.
- A 'proceed to show workings in workings area' button that triggers similar series of actions as that of 'proceed to elimination working area' button. This 'proceed to elimination working area' button is in Figure 6.20.
- A 'Set New unknown and value Computation Method' button produces similar actions as the 'Set New Variable and Elimination Method' button in Figure 6.20.
- A text area that displays the selected unknown and its value's computation method.
- A bigger text area to record details of the working that is based on the provided method.
- A 'Submit unknown's value computation workings' button to submit the produced working.

To compute for the value of unknown, the prototype needs a method to compute the value of unknown and the unknown object. If either input is missing and students try to proceed to the working area, the prototype responses with the message that is shown in Figure 6.26.



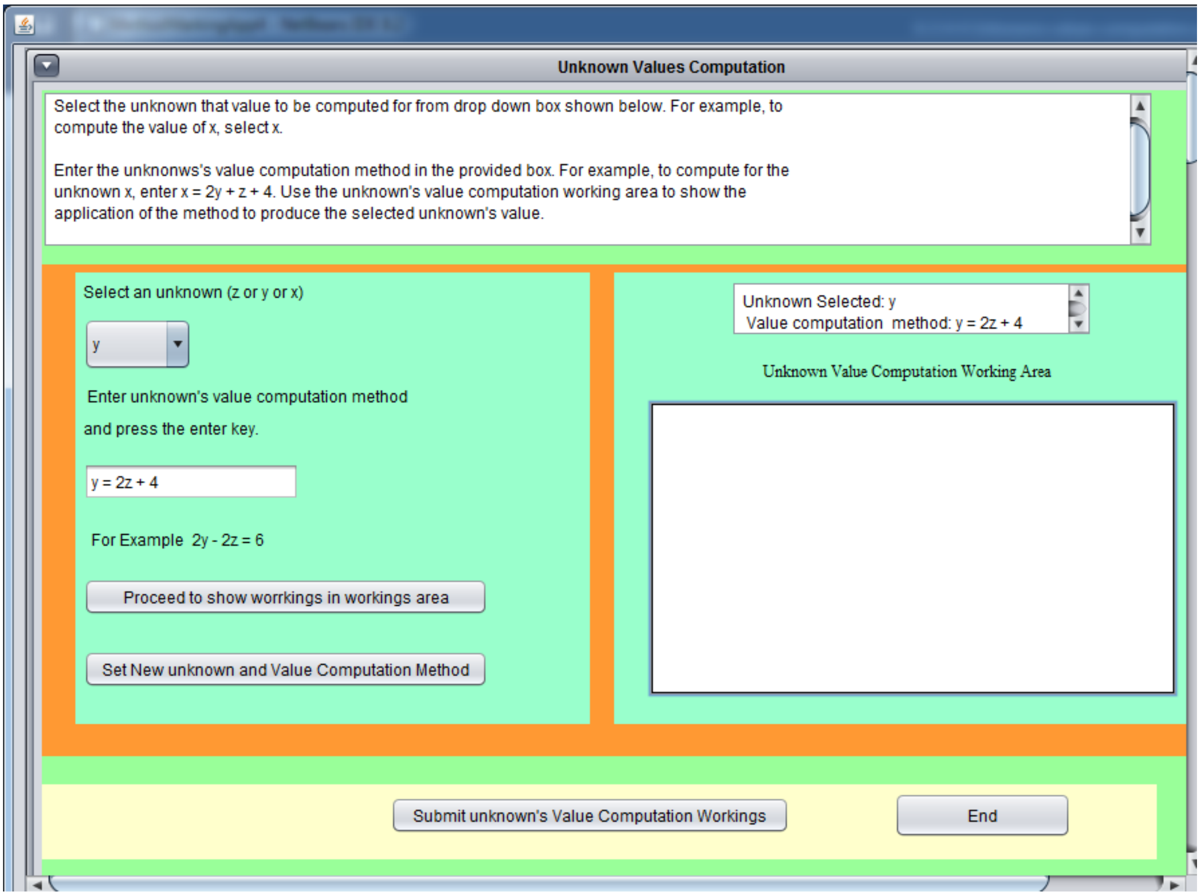
Select the unknown that value to be computed for from drop down box shown below. For example, to compute the value of x, select x.

Enter the unknonws's value computation method in the provided box. For example, to compute for the unknown x, enter x = 2y + z + 4. Use the unknown's value computation working area to show the application of the method to produce the selected unknown's value.

Please select the unknown that its value is to be computed for and enter the value computation method for the selected unknown. ==== The working area is only editable once you provide the two inputs.

Figure 6.26 Prompt to select unknown and its value computation method

When a student provides the two inputs, the prototype places the cursor in the working area. This scenario is shown in Figure 6.27.



Select the unknown that value to be computed for from drop down box shown below. For example, to compute the value of x, select x.

Enter the unknonws's value computation method in the provided box. For example, to compute for the unknown x, enter x = 2y + z + 4. Use the unknown's value computation working area to show the application of the method to produce the selected unknown's value.

Select an unknown (z or y or x)

y

Enter unknown's value computation method and press the enter key.

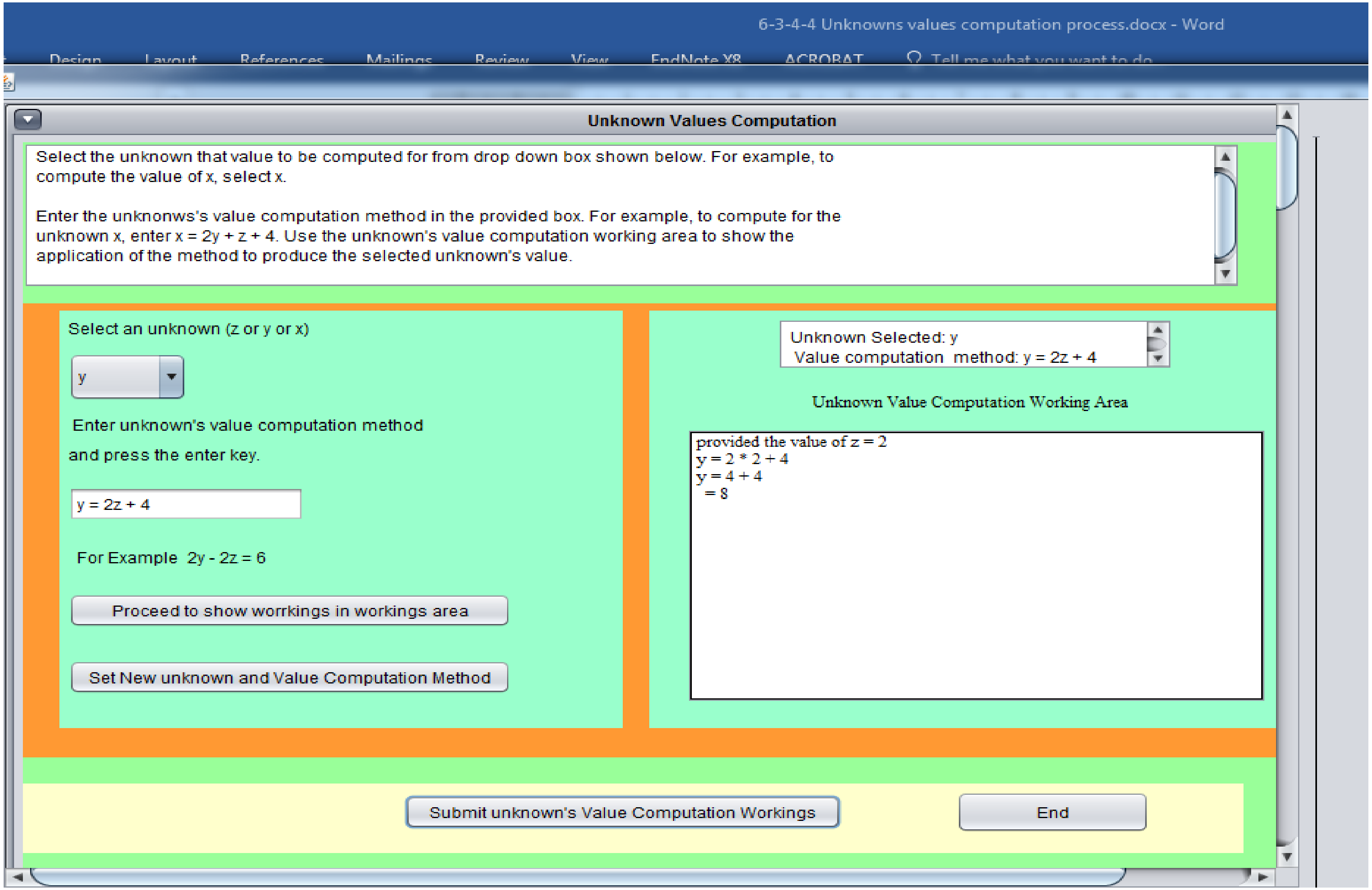
y = 2z + 4

For Example 2y - 2z = 6

Unknown Selected: y
Value computation method: y = 2z + 4

Figure 6.27 Unknown and its value computation method

Figure 6.28 shows a student working that has been entered by the student and saved in a file. The result of the evaluation of the code that is used to allow the student to compute the value of unknown and the explanation of the result are shown in Table 6.18. *All the messages that are generated and displayed in the console are for the purpose of testing the functionalities of the prototype. They are not sent to students. A communication between students and the prototype is achieved only through user interfaces.* The result of the evaluation and the observations that have been made during the value computation process of this prototype indicate that the prototype has performed as expected.



Select the unknown that value to be computed for from drop down box shown below. For example, to compute the value of x, select x.

Enter the unknonws's value computation method in the provided box. For example, to compute for the unknown x, enter x = 2y + z + 4. Use the unknown's value computation working area to show the application of the method to produce the selected unknown's value.

Select an unknown (z or y or x)

y

Enter unknown's value computation method

and press the enter key.

y = 2z + 4

For Example 2y - 2z = 6

Unknown Selected: y
Value computation method: y = 2z + 4

Unknown Value Computation Working Area

provided the value of z = 2
y = 2 * 2 + 4
y = 4 + 4
= 8

Figure 6.28 Unknown value computation process

Table 6.18 Unknown value computation process and outcome

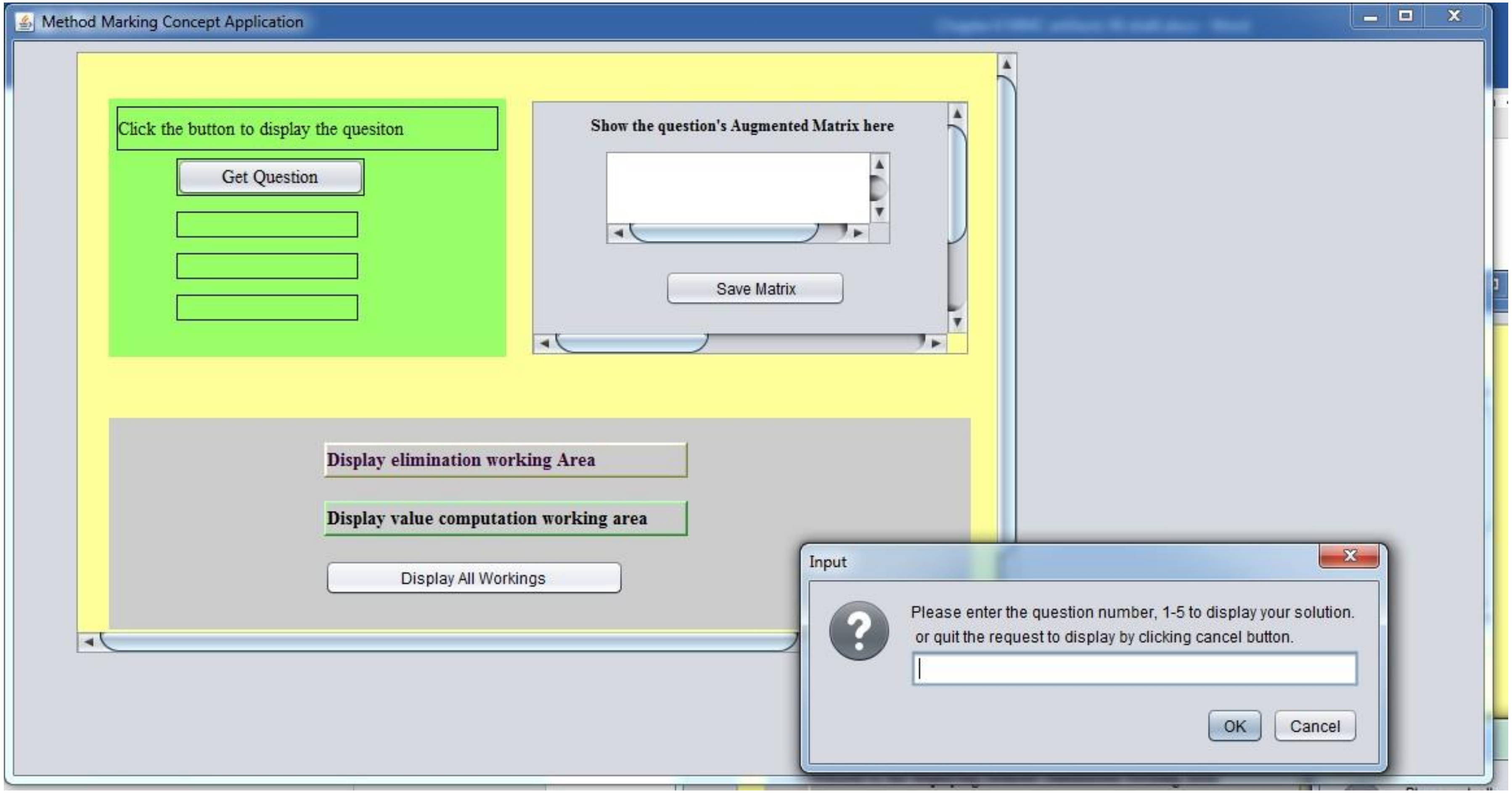
Messages that have been printed during 'unknown value computation' process and the process outcome	Explanations
run: Test: the question number is: 3 ======= Test: The question displayed	Stage 1 At this stage, the question to be solved is the question 3 and this question has been displayed.
Test: the unknown is selected.	The unknown whose value to be computed has been selected
Test: The working area is empty.	The student trying to submit an empty working
Test: work contains: provided the value of z = 2 y = 2 * 2 + 4 y = 4 + 4 = 8	The working including the unknown and the method that has been used to compute the value.
Test: The unknown value computation work is going to be written in the file.	A message that is being sent to the student that his/her working is going to be saved. This message is sent to the student using a dialog box user interface.
Test: the elimination works data to be written in the file provided the value of z = 2 y = 2 * 2 + 4 y = 4 + 4 = 8	These are workings that have been entered by the student and will be saved in a selected file.
Test: the option is yes	The student accepted to submit, and the system has saved the workings.
file .\src\problemSolversWorksFiles\q3\yvarValCompWork.txt created.	The student workings are saved in q3 directory which is the directory that contains the workings that belong to question 3. The 'yvarValCompWork.txt' is the file name that contains the working.
Test: Your work is successfully written in the file.	This message that shows that the file has been successfully saved. This same message has
Test: You opted not to compute for any other variable's value.	This message shows the student has chosen not to proceed with the process of computing a value for another unknown. Again, this
The messages in this table are not meant for students. They report the results of the evaluation of tasks in this part of this prototype.	

Sections 6.3.4.2, 6.3.4.3 and 6.3.4.4 have shown the functions that allow students to prepare their working and submit them. Section 6.3.4.5 allows students to display all or some of their workings and gives them options to update or delete.

6.3.4.5 Displaying and updating student working.

The tasks that will be performed by this part of this prototype will be divided into three sub-tasks. The first sub-task deals with actions related to requesting the display of student workings. The second one is concerned with actions to be performed after receiving a valid request. The third sub-task concerns actions to be performed after displaying the workings.

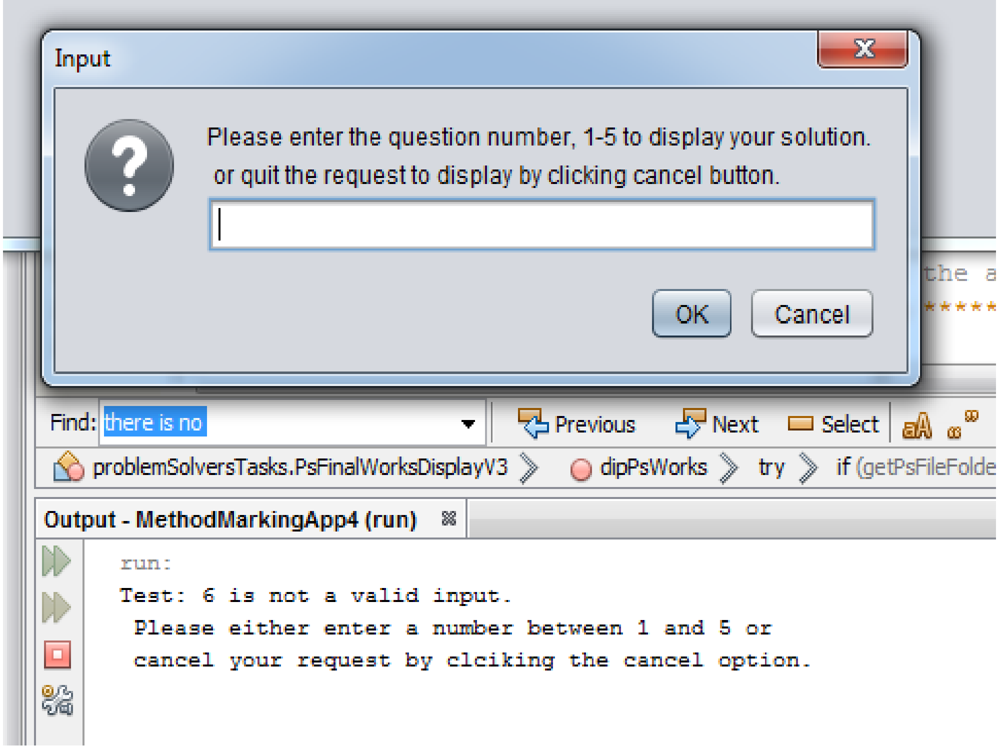
The first sub-task is initiated when a student clicks the 'Display All workings' button. There are multiple scenarios which are based on the input the student enters in the dialog box that is shown in Figure 6.29. The dialog box asks the student to enter a number between 1 and 5. If an invalid input is entered by the student, the dialog box remains displayed until the student enters a valid number or clicks the cancel button or closes the dialog box.



Please enter the question number, 1-5 to display your solution.
or quit the request to display by clicking cancel button.

Figure 6.29 Request to enter a question number to display workings

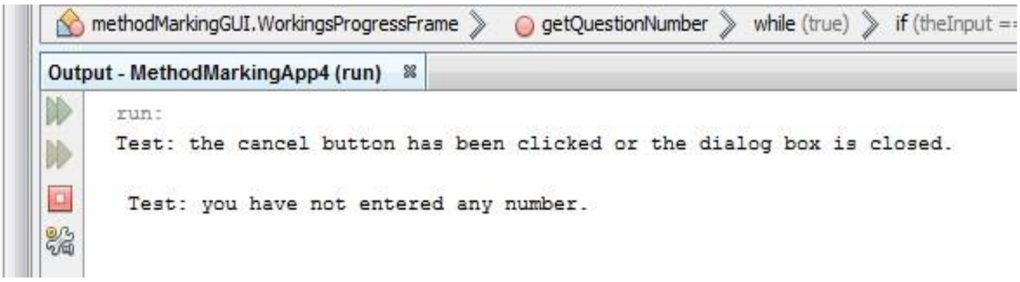
The results of providing an invalid input or clicking the cancel button or closing the dialog box are reported by the messages that are displayed in Figure 6.30 and Figure 6.31.



Please enter the question number, 1-5 to display your solution.
or quit the request to display by clicking cancel button.

```
run:
Test: 6 is not a valid input.
 Please either enter a number between 1 and 5 or
 cancel your request by clciking the cancel option.
```

Figure 6.30 Result of searching using invalid input



```
run:
Test: the cancel button has been clicked or the dialog box is closed.

 Test: you have not entered any number.
```

Figure 6.31 Result of clicking the cancel button or closing the dialog box

However, if the student enters a valid number, this prototype takes different actions. These actions will be addressed in this second sub-task.

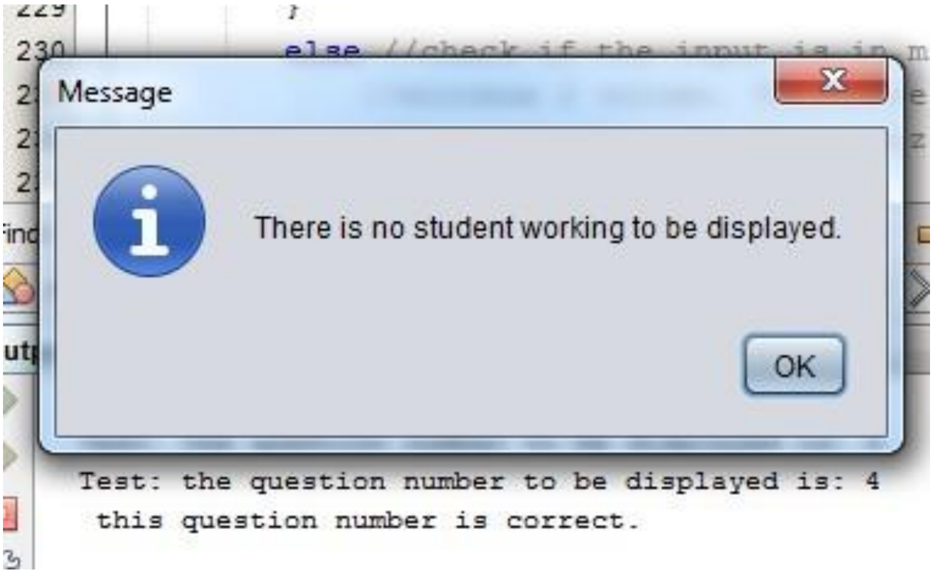
The discussion of the scenario of saving student workings helps to understand what this prototype does when the request to display a student workings is made. When a student working is saved in a file, the prototype uses the question number to search for the directory that will be used to contain the file that will be used to save the working. Files are only created if there are workings to be saved. However, the directories that contain these files have been created in advance. The number of directories that contains student workings are equivalent to the number of questions that require student workings. In total there are five directories and five questions. When a student asks this prototype to display his/her workings, the number that is provided by the student is used to check for the existence of the workings. The prototype requires the following components to display the requested workings.

- A scrollable container. This container is required due to the unknown number of dynamic panels.
- A main panel that is used as the container for the dynamic panels.
- Dynamic panels that contain input and output components.
- A main frame (window) that will be used to display the scrollable container.

The prototype performs the following actions:

- Create a directory name using the question number that is provided by the student.
- Use this directory name to create a file object that points to this directory. This file object is used to point to the directory that contains the files containing student workings.
- Create a File array using the file object.

At this stage all the components that have been listed above have been created except the dynamic panels and their components. The prototype uses the file object to check for the directory that contains the student workings. If this directory is empty, the prototype displays a message that is shown in Figure 6.32.



There is no student working to be displayed.

```
Test: the question number to be displayed is: 4
 this question number is correct.
```

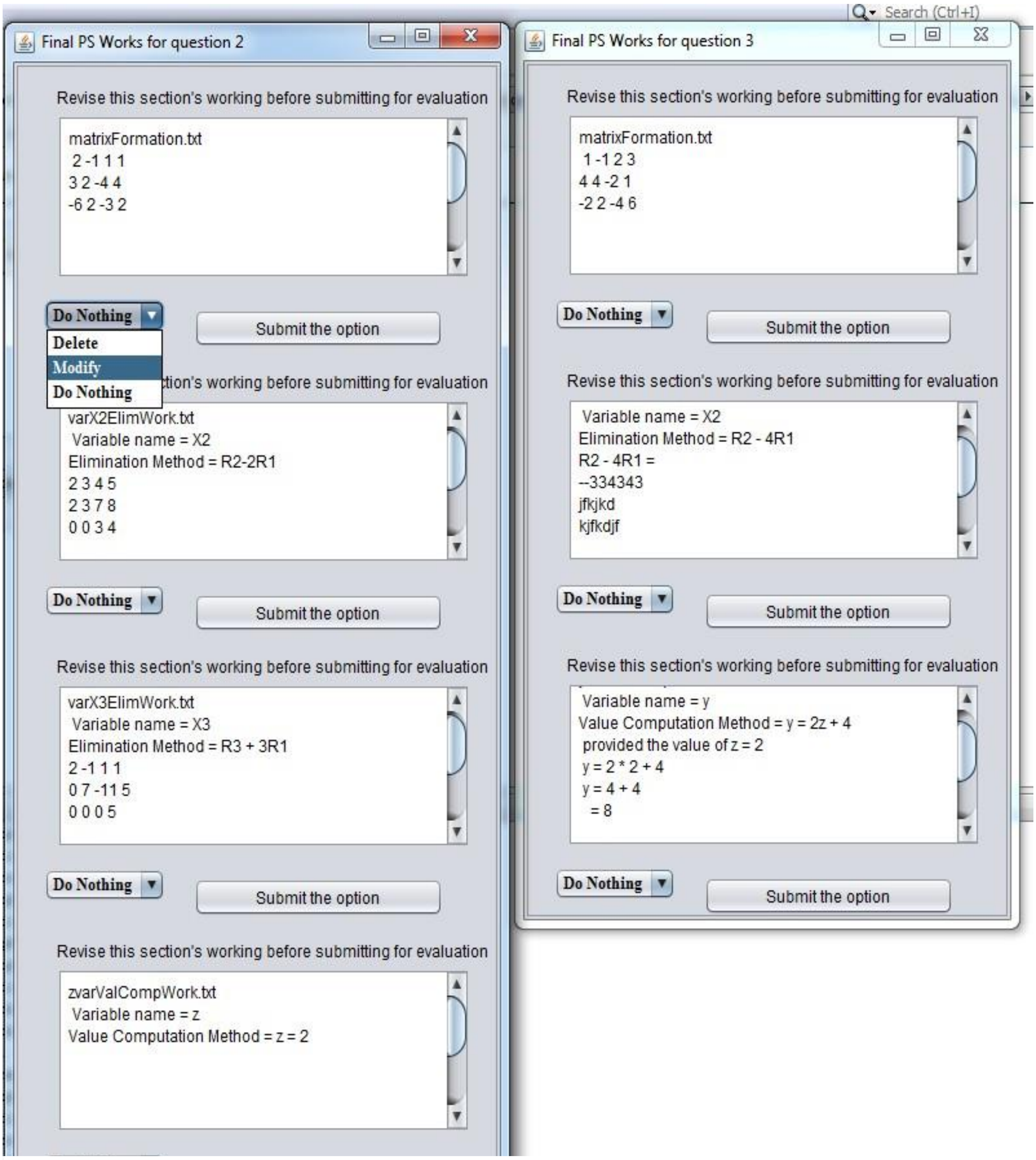
Figure 6.32 Result of searching for file in exiting directory

If a file exists or files exist, the prototype takes the following series of steps to display the workings.

- Prepares the main panel by setting its layout. The number of columns will be set to one while the numbers of rows are determined by the number of files that contain the student workings. As explained above the number of workings and the number of files is the same. For example, there may be four workings consisting of one matrix, and three element eliminations. The prototype saves this working in four separate files. Therefore, four separate dynamic panels are required to display these workings because each panel only displays one working. Conversely, the number of rows in the main panel is four.
- Creates a dynamic panel for each file in the array and displays the content of each file in the display area of the panel.
- Adds the dynamic panels to the main panel.
- Adds the main panel to the scrollable container.

- Adds this scrollable container to the frame (window).
- Displays the frame.

To demonstrate the above tasks, two questions with a varying number of workings are displayed in Figure 6.33. The number of workings displayed in the left window is four, while the number of workings in the right window is three.



Final PS Works for question 2:

```
matrixFormation.txt
 2 -1 1 1
3 2 -4 4
-6 2 -3 2
```

```
varX2ElimWork.txt
 Variable name = X2
Elimination Method = R2-2R1
2 3 4 5
2 3 7 8
0 0 3 4
```

```
varX3ElimWork.txt
 Variable name = X3
Elimination Method = R3 + 3R1
2 -1 1 1
0 7 -11 5
0 0 0 5
```

```
zvarValCompWork.txt
 Variable name = z
Value Computation Method = z = 2
```

Final PS Works for question 3:

```
matrixFormation.txt
 1 -1 2 3
4 4 -2 1
-2 2 -4 6
```

```
 Variable name = X2
Elimination Method = R2 - 4R1
R2 - 4R1 =
--334343
jfkjkd
kjfkdjf
```

```
 Variable name = y
Value Computation Method = y = 2z + 4
 provided the value of z = 2
 y = 2 * 2 + 4
 y = 4 + 4
   = 8
```

Figure 6.33 Student working from two questions

Table 6.19 shows the messages that have been created during the execution of the code that is used to display the workings using the questions number 2 and 3. These messages are generated at the result of testing the code at different stages when searching for a working to be displayed. The messages show two directory names that contain numbers 2 and 3. These numbers are provided by the student to search for q2 and q3 directories. The q2 directory contains four files, while q3 contains three files. These factors are demonstrated by the number of student workings that have been displayed in Figure 6.33.

Table 6.19 Result of searching for two different questions' student workings

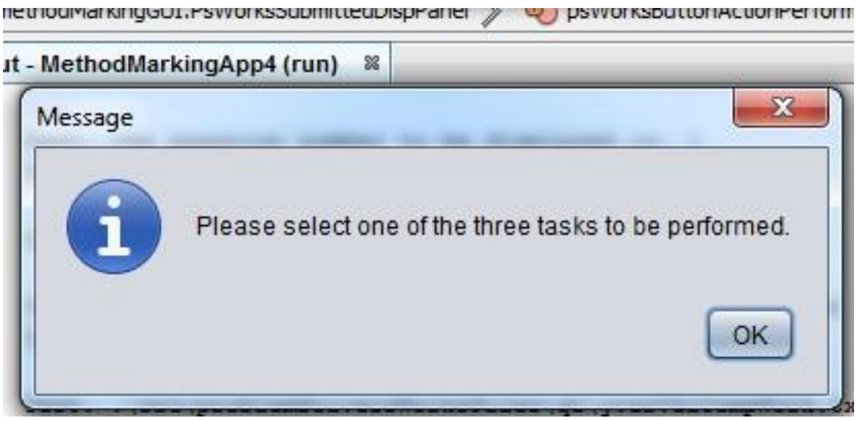
Results of searching for student workings	Explanations
Test: the question number to be displayed is: 2 this question number is correct. Test: found the file: matrixFormation.txt .\src\problemSolversWorksFiles\q2\matrixFormation.txt Test: found the file: varX2ElimWork.txt .\src\problemSolversWorksFiles\q2\varX2ElimWork.txt Test: found the file: varX3ElimWork.txt .\src\problemSolversWorksFiles\q2\varX3ElimWork.txt Test: found the file: zvarValCompWork.txt .\src\problemSolversWorksFiles\q2\zvarValCompWork.txt	The result of the evaluation shows that the number 2 that has been entered by the student is correct. Using this number, the prototype has found four files. Their names with directory they are placed in are shown.
Test: the question number to be displayed is: 3 this question number is correct. Test: found the file: matrixFormation.txt .\src\problemSolversWorksFiles\q3\matrixFormation.txt Test: found the file: varX2ElimWork.txt .\src\problemSolversWorksFiles\q3\varX2ElimWork.txt Test: found the file: yvarValCompWork.txt src\problemSolversWorksFiles\q3\yvarValCompWork.txt	This result of the evaluation reveals the same type of information as the one above. The differences are the files store workings that are submitted for question 3 and the number of workings is also three.

Once the student workings are displayed, the student can operate on the displayed workings. The third sub-task addresses actions that student can perform using the components that are contained in the dynamic panel. These components are listed below:

- A text area that is used as an input and output component. This area is used for displaying the student workings and it is also can be used as an input component.

- A drop-down list. The drop-down list is an input component that allows students to select one of the options which are ‘delete’, ‘modify’ and ‘do nothing’.
- A ‘Submit the option’ button to submit an option that is selected from the drop-down list. Each button on each dynamic panel is independent of the other buttons.

To perform a selected option from the drop-down list, the prototype checks for the clicking of the ‘Submit the option’ button. However, if the button is clicked without selecting an option, the prototype responds with the dialog box that is displayed in Figure 6.34.



Please select one of the three tasks to be performed.

Figure 6.34 Request for an option selection

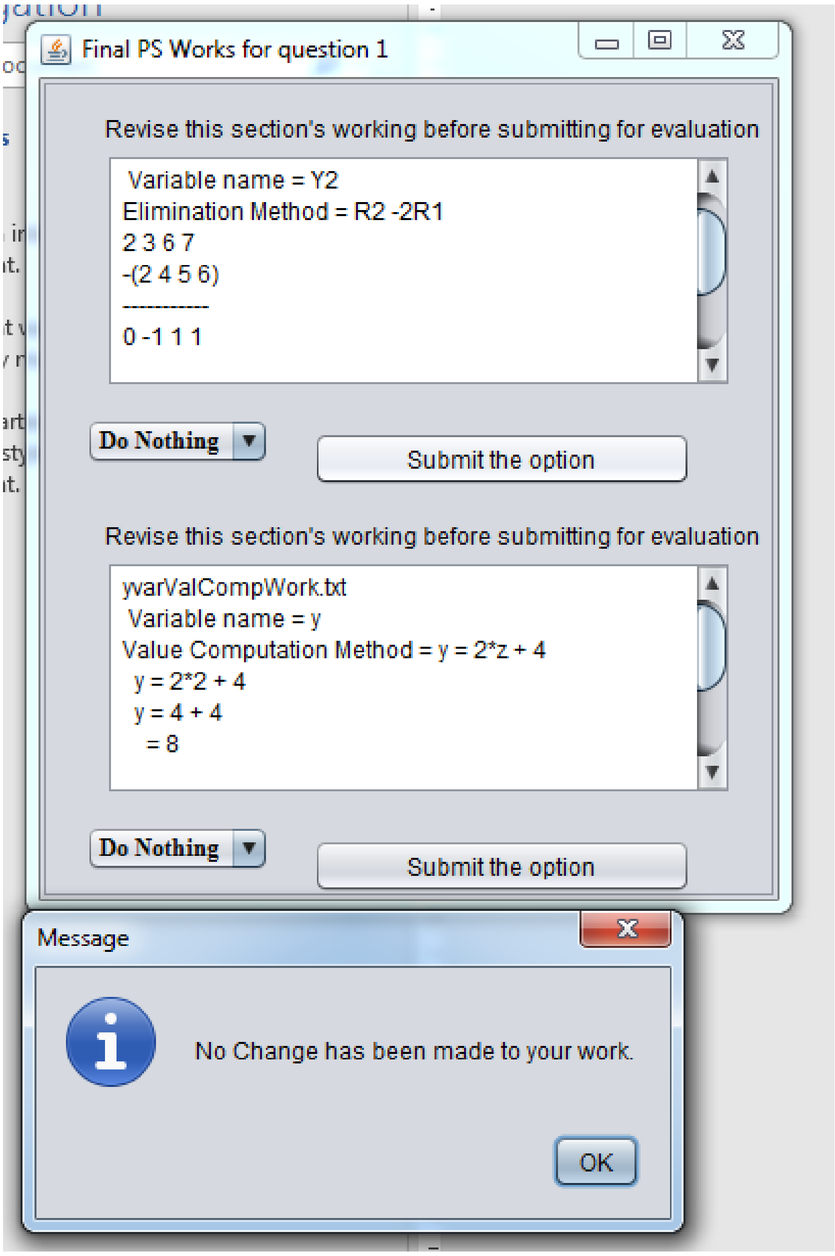
If the student selects one of the options, the prototype performs the following actions:

- Accesses the File array object that has been created to keep the files that contain the student workings.
- Reads a file name from the array object.
- Reads the student workings from the text area of this dynamic panel for which its button has been pressed.
- Checks if the file name that has been read from the array object can be found in the student workings.

- If the file name is found in the student workings, saves this file name.

- Checks the student option and do the following.

  - If the 'Do nothing' option is selected, the message in Figure 6.35 is displayed and the prototype continues displaying the student workings. Table 6.20 displays the result of the evaluation of the prototype for this part of functionality.



Revise this section's working before submitting for evaluation

```
 Variable name = Y2
Elimination Method = R2 -2R1
2 3 6 7
-(2 4 5 6)
-----------
0 -1 1 1
```

Revise this section's working before submitting for evaluation

```
yvarValCompWork.txt
 Variable name = y
Value Computation Method = y = 2*z + 4
  y = 2*2 + 4
  y = 4 + 4
    = 8
```

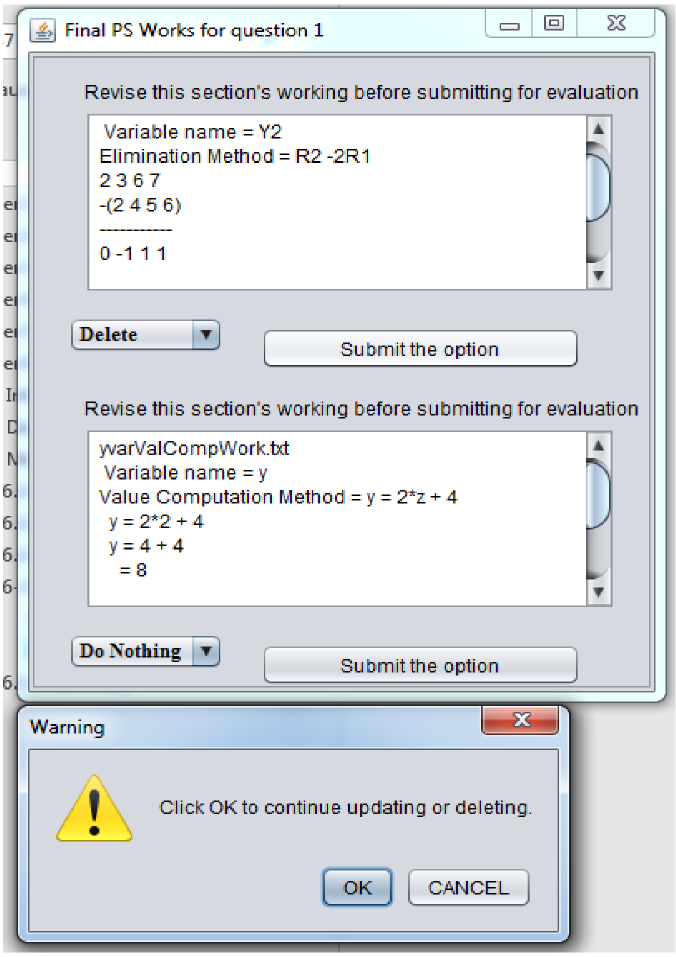
No Change has been made to your work.

Figure 6.35 Do nothing option selection

Table 6.20 Result of testing for 'Do nothing' option

Results of testing the ‘Do nothing’ option	Explanations
Test: this is from the student workings display area. varY2ElimWork.txt Variable name = Y2 Elimination Method = R2 - 2R1 2 3 6 7 -(2 4 5 6) ---------- 0 -1 1 1 Test: The ‘do nothing’ option has been selected:	The test code reads the content of the file after the ‘do nothing’ operation has been performed. The content of the file displays the same content as that is displayed in the display area of the dynamic panel. This shows nothing has been changed in the content of this working. At the end of this evaluation, a message is displayed to indicate that no change has been made to the original working. This message has also been sent to the student using the dialog box as shown in Figure 6.35.

- If the ‘delete’ option is selected, the prototype (1) deletes the file from the directory and (2) removes the dynamic panel from the main panel. However, the prototype displays a warning message before it performs these actions. Figure 6.36 shows the warning message to allow the student to confirm the deleting option or to cancel the delete option.



Revise this section's working before submitting for evaluation

```
 Variable name = Y2
Elimination Method = R2 -2R1
2 3 6 7
-(2 4 5 6)
-----------
0 -1 1 1
```

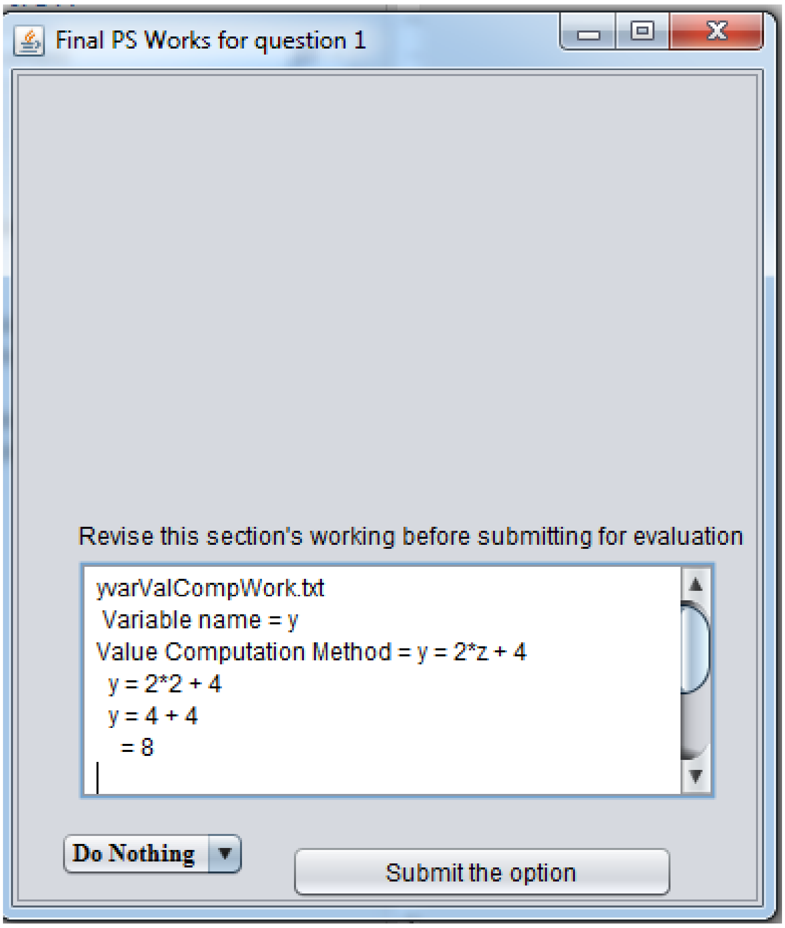
Revise this section's working before submitting for evaluation

```
yvarValCompWork.txt
 Variable name = y
Value Computation Method = y = 2*z + 4
  y = 2*2 + 4
  y = 4 + 4
    = 8
```

Click OK to continue updating or deleting.

Figure 6.36 Figure 6.36 Delete student working option

When the student clicks the ‘OK’ button, the prototype completes the two actions that have been mentioned above. Figure 6.37 shows the window that shows the removed workings area as blank.

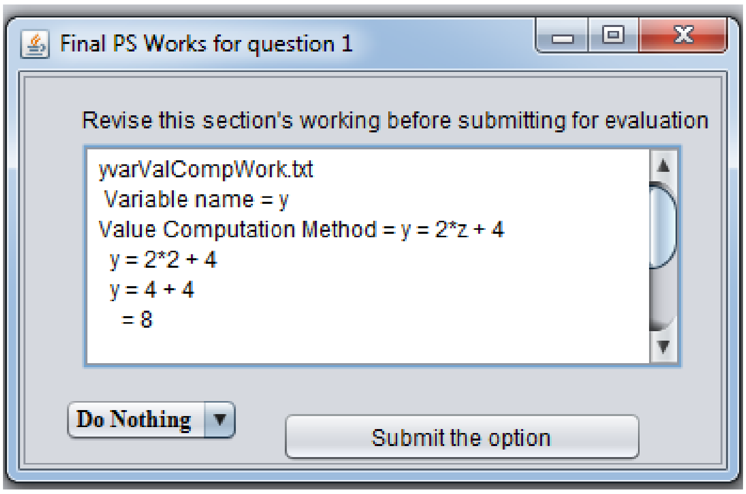


Revise this section's working before submitting for evaluation

```
yvarValCompWork.txt
 Variable name = y
Value Computation Method = y = 2*z + 4
  y = 2*2 + 4
  y = 4 + 4
    = 8
```

Figure 6.37 Result of deleting the requested working

To check that the deletion of the working is permanent, a new request can be made to redisplay the workings of question 1 to see the change that has been made. This change is shown in Figure 6.38 in which one working is only displayed. The comparison of the window in this figure with the window in Figure 6.37 reveals that both figures show only one dynamic panel.

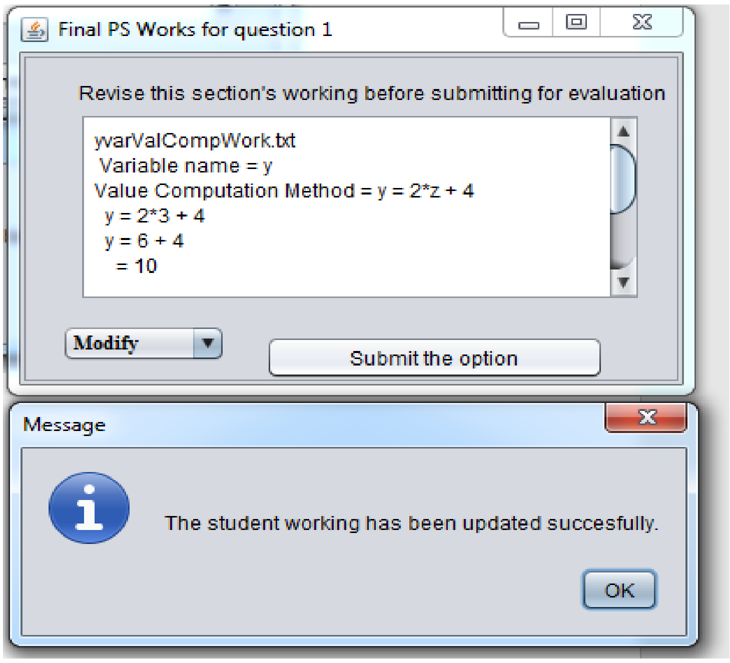


Revise this section's working before submitting for evaluation

```
yvarValCompWork.txt
 Variable name = y
Value Computation Method = y = 2*z + 4
  y = 2*2 + 4
  y = 4 + 4
    = 8
```

Figure 6.38 Result of requesting question 1 workings

- If the ‘modify’ option is selected, the prototype reads the modified data from the text area where the original student workings have been displayed. Then it checks for the availability of the file that was the source of the displayed working. In the next step the prototype writes the modified working in this file and displays a warning message and options to continue with the modification process or to cancel the ‘modify’ request. Figure 6.39 displays this situation.



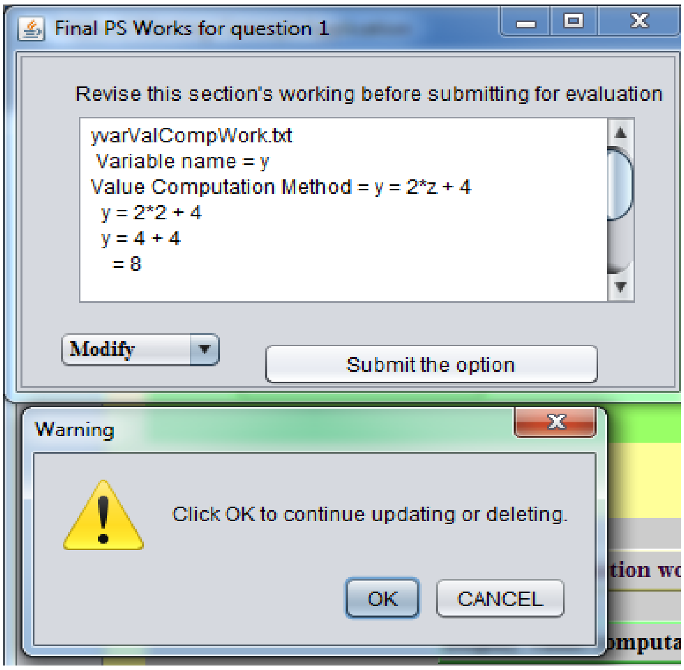
Revise this section's working before submitting for evaluation

yvarValCompWork.txt
Variable name = y
Value Computation Method = y = 2*z + 4
y = 2*3 + 4
y = 6 + 4
= 10

The student working has been updated succesfully.

Figure 6.39 Request of modification option

- However, when the student confirms the modification to be made, the prototype completes this action and displays a message that is shown in Figure 6.40.



Revise this section's working before submitting for evaluation

yvarValCompWork.txt
Variable name = y
Value Computation Method = y = 2*z + 4
y = 2*2 + 4
y = 4 + 4
= 8

Click OK to continue updating or deleting.

Figure 6.40 Request to confirm student workings modification or cancelling option

Table 6.21 displays the results of testing the code at different stages of the student workings modification process. Furthermore, this table gives the explanations of the message in the table.

Table 6.21 Test result and explanation of student working modification result

Test result of student workings modification process messages

```
Test: the question number to be displayed is: 1
This question number is correct.
Test: found the file: yvarValCompWork.txt
 .\src\problemSolversWorksFiles\q1\yvarValCompWork.txt
Test: .\src\problemSolversWorksFiles\q1\yvarValCompWork.txt

Test: this is from the student workings display area.
yvarValCompWork.txt
 Variable name = y
Value Computation Method = y = 2*z + 4
  y = 2*2 + 4
  y = 4 + 4
    = 8

Test: The selected option is: Modify
Test: The modify option has been made:

 Test: File name is to be written: yvarValCompWork.txt

Test: Your work has been successfully updated in the file.

Test: this is from the student workings display area.
yvarValCompWork.txt
 Variable name = y
Value Computation Method = y = 2*z + 4
  y = 2*3 + 4
  y = 6 + 4
    = 10
```

explanation	The testing confirms that the requested question is 1 and it is a valid number. This number was used to form the name of the directory q1. The prototype used this directory name to search for the existence of this directory. The message in the table confirms the existence of this directory. The prototype continued the process by reading the name of the file from the directory q1 and searched this file name in the student workings and found this file's name on the first line of the workings. The 'found the file: yvarValCompWork.txt' message in the table confirms the existence of the file in this directory. The table shows the original content of the student workings containing the file name. The prototype continued the process by reading the modified student workings from the display area of the dynamic panel and saving new content in the file. The table shows these new modified workings that are read from the file. These test messages indicate the success of the evaluation in testing the prototype to achieve the purpose of modifying student workings.

6.4 Summary

The purpose of this chapter was to develop artefacts of this DSR. To achieve this purpose, this chapter began the discussion about types of DSR artefacts that are accepted as DSR project outputs. The discussion has revealed characteristics of these types of artefacts. These include solving significant complex problems and/or making significant improvement in existing artefacts that are used in solving these problems. Furthermore, the discussion has found the existence of various types of artefacts as DSR outputs. There are varying views about types and numbers of these artefacts. The major outputs that have been discussed in this chapter were 'constructs', 'models', 'methods' and 'instantiations'. The decision of what types of artefacts this DSR produce has also been made. In this research the focus was on building constructs and models. However, the methods artefact and a prototype of the instantiation artefact were also discussed and created. However, before discussing and building these artefacts, the chapter discussed the importance of 'design', 'design science' and 'design science research' terms.

The meanings of the term 'design' as a noun and a verb have been discussed. The usage of existing knowledge or producing new knowledge to generate an implementable plan, as the output of a design process, has been recognised. The dependency of creating an innovative design using design science as an approach has also been understood. The role of DSR in producing the knowledge that is required by design science has also been established. In this research, this knowledge comes from the efforts that have made in different stages of this research. The process of creating this knowledge and the knowledge itself have been discussed in the previous chapters.

The first artefact that this chapter discussed and constructed was the constructs. The constructs were composed after their sources were identified. The main contributors were the analyses of student scripts (investigated context), kernel theories, the outcome the literature review and justificatory knowledge. As defined in section 6.3.1, constructs artefact is the foundation of other artefacts. The second artefact that was created was models.

The models were created using constructs and the sources that have been used to create constructs. Models were defined as the relationships between constructs and are used in explaining problems and producing solutions. The relationships between constructs were analysed to form these models. This chapter showed how the analyses of scripts were used to construct models that reflect students' actions when producing their workings. One example of the workings is the solution strategies used in producing answers. Other models were also constructed based on the knowledge that was extracted from the sources that have been mentioned above. Moreover, these models showed tasks that assessors and students can perform when these models are implemented in an instantiation artefact.

The chapter continued to the next step which was defining a DSR methods artefact and developing it. The roles of methods in implementing models were identified. The existence of different strategies to create DSR methods was explained and the flowchart technique has been selected. The flowcharts were presented and steps in these flowcharts were implemented when they were used to create one of the functionalities of the prototype.

In the creation of the prototype five questions were stored in a file. The prototype allows a student to request any of these questions and perform these activities: form a matrix, eliminate elements, compute values of unknown, retrieve saved workings and display them. Additionally, the prototype gives options of deleting, or updating the workings.

This chapter also addressed the issue of evaluating the functionalities of this prototype. Testing code was embedded in selected parts of the prototype to report the status and responses of the prototype. The prototype and students' actions were monitored and reported in console. These reports or messages were analysed and compared with the observations of the prototype functionalities. The observations and messages approaches were adopted as the evaluation method in testing the prototype. The results of the evaluation have shown that the prototype performed as designed. The user interface of the prototype also functioned as designed in allowing students to perform the tasks that have been outlined above. Many iterations were made to improve the efficiency and functionalities of the prototype. Thus, this chapter has added design knowledge that can be claimed as significant knowledge contribution.

Chapter 7 Evaluation

Previous chapters have contributed to the development of DSR artefacts in this research. These artefacts have been evaluated at different phases of their creation. The creation and evaluation of the main constructs have been discussed in chapter five. The main models were also created and theorised in chapter five. Chapter six discussed and created constructs, models, methods and prototype artefacts. The evaluation of the methods and prototype artefacts has been performed in chapter six.

The purposes of this chapter are to discuss the importance of evaluation in a DSR, to explain the purposes of evaluation, discuss evaluation methods and to identify attributes of this DSR's artefacts.

This chapter will focus on the following topics:

- Definitions of evaluation, validation and the relationships between them.
- Importance of evaluation.
- Different purposes of evaluation and selecting appropriate purposes for evaluating this DSR's artefacts.
- Scope of evaluation in this research.
- Different evaluation methods in DSR and methods that have been selected for evaluating this DSR's artefacts.
- Relationships between build and evaluate activities.
- Prototype as evaluation method and as an artefact.
- Attributes of this project's artefacts.
- Roles of feedback in an evaluation process.

When constructing the DSR framework for this research, selected published DSR frameworks were reviewed. The review of the evaluation process in the selected frameworks has been shown in section 3.2.4. This chapter is the continuation and the expansion of that section. Different definitions are given to the terms of ‘evaluation’ and ‘validation’. For example, Gettelman and Rood (2016) define an evaluation of a model as means to reveal the efficiency of a model when performing a task. On the other hand, Marwedel (2011) describes an evaluation as a process that involves “computing quantitative information of some key characteristics (or “objectives”) of a certain (possibly partial) design” (p. 204). Gettelman and Rood define a validation as the process that involves examining the veracity or accuracy of an artefact when performing a specified task. Validation is also defined as a process that examines the relevancy of a design to show that “a certain (possibly partial) design is appropriate for its purpose, meets all constraints and will perform as expected” (Marwedel, 2011, p. 203). The validity of a design is established by the result of the evaluation of artefacts confirming the purpose of the design (Pries-Heje & Baskerville, 2008). The definitions of these two terms (evaluation and validation) lead to the understandings that an evaluation examines the value of an artefact while the validation examines the extent of the ‘truth’ in the claimed value. Hence, evaluation and validation are important components or activities in DSR and in all research generally. In this research, both activities have been performed to achieve the purpose of this research.

The outputs of both DSR and IS research projects require evaluation (Gregor & Jones, 2007; March & Smith, 1995; Venable, Pries-Heje, & Baskerville, 2016). Venable et al. (2016) report that there are six purposes for performing evaluation in DSR. Two of these purposes are shown below.

- “First, one key purpose of evaluation in DSR is to determine how well a designed artefact or ensemble of artefacts achieves its expected environmental utility (an artefact's main purpose)" (p. 79).

- "A fourth purpose considers that utility is a complex, composite concept, which is composed of a number of different criteria beyond simple achievement of an artefact's main purpose" (p. 79).

At the start of this research, the main aim was to create an artefact that assesses subjective questions. However, the complexity of this issue became apparent when the outcome of the literature review and the analyses of the student scripts were examined. The artefacts that have been created in this research can be evaluated for the first purpose provided that the main purpose is contextualised to the domain knowledge that student scripts examine. To evaluate the artefacts for the fourth purpose the main purpose of assessing subjective questions must be taken into consideration. In this case, a simple achievement is expected from the artefacts. Therefore, the two purposes that have been mentioned above need to be taken into consideration when examining the results of the evaluations of the artefacts that have been produced in this research. After determining the purpose of evaluating artefacts, it is essential to specify what to evaluate.

In DSR and other related research, the evaluation process examines artefacts in terms of their "functionality, completeness, consistency, accuracy, performance, reliability, usability, fit with the organization, and other relevant quality attributes" (Hevner et al., 2004, p. 85). According to Hevner et al., in DSR, "the principal aim is to determine how well an artefact works" (p. 88). However, the scope of evaluation process is not limited to DSR artefacts. It encompasses "the rigorous application of research methods and the robustness of research results" (Szopinski, Schoormann, & Kundisch, 2019, p. 4). In other words, "it must address the project's broader research questions by validating adopted theory or leading to the generation of new theory" (Prestopnik & Crowston, 2012, p. 76). Therefore, the purposes of evaluation are not only concerned with ensuring the production of a purposeful artefact but also includes the concern of adding new knowledge (Venable et al., 2016). Thus, this includes the evaluation of the entire research.

Different sections of this research have been evaluated at different phases. After determining the research purpose, the context (student scripts) to be analysed had been evaluated for its relevance for this research purpose. Research philosophies were investigated and evaluated to help determine the most relevant methodologies and methods. The selected methodologies and methods were also investigated and evaluated. Then these methods were used to analyse student scripts, conduct the literature review and build artefacts. These activities have been performed in chapters two, three and four. The reliability and validity of the analysis process of the scripts and the outcome of the analysis have been performed in section 5.1. The credibility of the entire activity has also been established in the same section. In section 5.2 the findings in the qualitative data have been conceptualised and used as fundamental knowledge to form constructs and models that have been created in chapter six. The iteration of the analysis process of scripts has been shown in Figure 5.1. Furthermore, the iteration process involving build, evaluate and feedback tasks has been performed when creating this research DSR artefacts.

The above discussion is one of the methods that have been used to evaluate constructs and models in this research. The process that was used to create the questionnaire to conduct the quantitative analysis of the scripts was another evaluation method that refined qualitative data. Moreover, the quantitative analyses of the scripts added another type of evaluation that included the refining of the qualitative analysis data.

These types of evaluation methods are appropriate for evaluating artefacts in a DSR and IS research projects. Although one of the evaluation definitions that has been given above considers evaluation methodology as quantitative, there are other authors, such as Venable et al. (2016) and Herselman and Botha (2015), who also accept the interpretive paradigm that uses qualitative methodology. The evaluation methods that have been explained above are qualitative. For further evaluation of the analyses of scripts, some of the findings of that analyses have been compared with the previously published findings. This example has been shown in section 5.4.2.3. This form of evaluation is an example of the 'logical argument' method which is one of the eight types that have been classified by Peffers, Rothenberger, Tuunanen, and Vaezi (2012). The 'prototype' evaluation method is also one of the eight evaluation methods that are classified by Peffers et al. All the above- mentioned types of evaluation methods and evaluation processes increases the acceptability of constructs and models artefacts.

The evaluation methods that have been used to evaluate methods artefact and prototype artefact are testing and observation. The methods artefact has been tested at different levels. The first testing involved going through steps in flowcharts to check if they behaved as expected. Section 6.3.3 demonstrated the creation and implementation of methods. Testing, using debugging techniques, and observations were the two methods that have been used to evaluate the prototype artefact. The technique of embedding debugging code in the prototype code that performs the required tasks shows the close relationships between building and evaluating DSR artefacts. These relationships are stated by Venable et al. (2016) as follows:

> As part of the design science process, evaluation may be tightly coupled with design itself. This tight linkage arises from the impact of evaluations on designer thinking, with the potentially rapid cycles of build and evaluate that sometimes constitute design itself. (p. 78)

As explained above, the prototype artefact is one of artefacts evaluation methods. In addition, the prototype is evaluated as artefact. In this research, the prototype has implemented some of the utilities that constructs artefact and models artefact provide.

The attributes of this research artefacts are shown in Table 7.1. Methods and prototype artefacts partially implemented the functions in the models. To show this status, the methods and prototype attributes are shown using dots.

Table 7.1 MMC artefacts attributes

	Constructs	Models	Methods	Prototype
• Functionalities • Consistency • Accuracy • Performance • Reliability • Usability • Fit with the organisation	✓	✓	●	●
Completeness				

The functionalities of these artefacts have been discussed in sections 6.3.1, 6.3.2, 6.3.3 and 6.3.4. The evaluation and validation processes that have been explained above ensured that the constructs and models artefacts have achieved the attributes that are shown using ticks in the table. The tick is used to indicate that these attributes cover all the constructs that have been created as well as all the features in the models' artefact. The 'Completeness' attribute is not assigned to any artefacts. There might be limitation to claim the 'Completeness' attribute for constructs and models. Some of the reasons are (1) partial implementation of the models in the prototype and (2) difficulties in understanding student errors affecting the creation of some constructs. The complexity of understanding student workings has been discussed in section 5.2. The attributes that have been assigned to methods and prototype artefacts are limited to the parts that have been implemented only. Therefore, for both the absence of the completeness attribute is understandable. The quality of these attributes can be improved through the 'build and evaluate' iteration.

The nature of iteration in 'build and evaluate' involves feedback, which modifies the design process and design product (Hevner et al., 2004) in a DSR project. The nature of the evaluation activity is not only to complete a research phase, but it is also to propose ideas that enhance the artefacts (Prestopnik & Crowston, 2012). The evaluation processes that have been used in this research have contributed to the production of knowledge that will be claimed in the discussion chapter.

This chapter has identified the evaluation and validation terms and their relationships. The importance of evaluating DSR artefacts to claim their validity has been established. Different purposes of evaluating DSR artefacts have been explained. Two purposes that have been identified to evaluate this research's artefacts were to check how well these artefacts achieved the purpose of this research and to the wider field. The discussion of the evaluation processes has revealed the scope of the evaluation, which covers the entire research activity.

The usage of different evaluation methods to evaluate and validate constructs artefact and models artefact have been discussed. These evaluation methods were the processes that have been used to select the appropriate methods of analysis, applying those methods to the student scripts, and conducting the literature review. The discussion has also included the methods that have been used to validate the analysis process to analyse student scripts. The legitimacy of using these evaluation methods when producing and validating constructs artefact and model artefact has been explained in this chapter.

The evaluation methods that have been used to evaluate methods and prototype artefacts have also been explained. The double roles of the prototype as an evaluation method and as an artefact to be evaluated has been discussed. The attributes of this DSR's artefacts in terms of utilities have been explained. The chapter ended by discussing the impact of feedback, in the evaluation processes, relating to the design process and design products in DSR in general.

Chapter 8 Discussion and conclusion

In previous chapter, the evaluation processes of this DSR's artefacts has been discussed.

The purposes of this chapter are to (1) summarise this research, (2) discuss the findings in this research, (3) identify the knowledge to be contributed, (4) explain the limitations in this research, (5) identify the research findings implication for practice and (6) to explain the future directions of this research.

Section 8.1 will give the summary of this research. Section 8.2 will discuss findings from the literature review and the analyses of the student workings. The complexities of assessing subjective questions and the need to produce the knowledge and ability to employ this knowledge in solving this complex problem will be discussed. The discussion will also include how the knowledge in these findings will be used to improve the assessment of subjective questions using DSR artefacts.

Furthermore, the lack of using the research philosophies in previous development of e-assessment systems will be discussed. This finding in relation to the development and utilities of both model artefact and prototype artefact will also be discussed. Additionally, the user interfaces design in the prototype in terms of contribution to this research purpose will be discussed.

The sections of this research that contributed knowledge will be discussed in section 8.3. The claimed knowledge will include the inference of solution strategies from the student scripts, the DSR artefacts which are constructs, models and prototype as well as the questionnaire instrument. The practical usage of the contributed knowledge will be identified. Furthermore, this section will identify the roles of researchers and of the knowledge base as the subjects that determine the sections that contribute the knowledge. Moreover, this section will explain knowledge acquisition methods that assist in acquiring knowledge from different sections in this project.

This research limitations will be identified in section 8.4. The numbers of e-assessment systems to be examined, the exclusion of some student scripts from the analyses, the possible occurrence of bias during the qualitative analysis of scripts will be among those limitations that will be discussed. Additionally, the limitations in both the DSR artefacts and their evaluation methods will be discussed. The future directions or recommendations of this research will be presented in section 8.5. These directions will be discussed in terms of the knowledge contribution of this research.

8.1 Summary of this research

Previous research publications in e-assessment of mathematics have generally been concerned with assessing objective assessments. However, this “is not enough to assess the student thoroughly” (Kumaran & Sankar, 2015, p. 24). These publications also found the importance of using mal-rules and common errors in improving the e-assessment for objective assessments. Greenhow (2015) elaborates that “algorithms for wrong answers (displayed in a multi-choice MC question, and behind the scenes for a responsive numerical input (RNI) question) should be encoded based on common mistakes or mal-rules” (p. 118). The purposes of this research were to understand why the e-assessments for mathematics face challenges to assess subjective questions and to produce the knowledge that helps understanding the difficulty and produce solutions. To understand this problem the literature review was conducted, and the output of this literature review showed the need for further study; the review indicated the advantage of using design science research (DSR) as the paradigm to conduct this research. Student scripts containing solutions to multi-steps questions (MSQ) was selected. In this research, the word scripts have been used interchangeably with workings that contain solutions of MSQ.

The purposes of analysing the scripts were not only to understand the difficulty (of e-assessment of subjective questions), but also to use the knowledge that would be created from the analyses to build and evaluate DSR artefacts. Therefore, the need of DSR for the knowledge that is produced by other sciences (natural and social sciences) research and their methodologies were discussed and established. The knowledge that will be produced by the analyses of the scripts is an example of this knowledge.

However, before starting the analysing of these scripts, multiple existing DSR frameworks were examined and a relevant DSR framework was created. Using this DSR framework as a guide, the research continued to the next stage which was the exploration of the research philosophies. Several different paradigms were investigated in terms of ontology and epistemology. The investigation of these philosophies resulted in selecting critical realism as the most appropriate paradigm to conduct this research. This paradigm was used to investigate the most relevant methodologies and methods to analyse the scripts, and to build and evaluate DSR artefacts corresponding to this research.

The selected methodology was mixed methods. In first phase, qualitative inductive content analysis method was used to analyse 258 student scripts using the following questions.

1) What types of solution strategies have been used to solve problems?
2) What types of errors have students made during solving problems?

The answers to these qualitative questions have been given using the information in subcategories of solution strategies and error of categories –these are student error- that have been created from the findings of qualitative analysis of the scripts. Figure 6.18 displayed solution subcategories while Table 5.1 and Table 5.2 explained the meanings of these subcategories. Figure 6.19 displayed subcategories of errors while Table 5.3 contained the meanings of these subcategories. The discussions about these qualitative findings, the process of creating categories and subcategories and challenges in the analysis, evaluation and validation of the qualitative analysis process and the output of the process were given in sections 5.4.1 and 5.4.2 and in appendix B. In total there were 12 types of solution strategies and 11 types of student errors that were found in the qualitative analysis of scripts.

In the second phase, the scripts were analysed quantitatively. Using the result of the qualitative analysis of the scripts, a questionnaire was made to collect data from 160 scripts and analysed statistically. The following questions were used to analyse the collected quantified data.

1) How many students have eliminated the fourth, seventh and eighth elements sequentially in 3x3 matrix (as show in Figure 8.1) to create an upper triangular matrix in a Gaussian elimination?
2) What types of solution strategies have been used to eliminate the fourth, the seventh and the eighth elements from the matrix?
3) What is the frequency of each type of error across all scripts and where, in the student working, does this error appear?

$$\begin{bmatrix} 1, 2, 3 \\ 4, 5, 6 \\ 7, 8, 9 \end{bmatrix}$$

Figure 8.1 Numeral representation of elements positions

The quantitative research questions answers were given in sections 5.5. These questions were answered during the presentation and discussion of the result of the quantitative analysis of the selected scripts.

The qualitative analysis of student scripts has involved interpretation and discussion during and after the analysis process. Additionally, the quantitative analysis of the scripts has also involved the presentation as well as the discussion of the quantitative findings. The purposes of the interpretation and discussion in both phases were to understand student workings deeply, so meaningful knowledge can be deduced. The meanings of the findings in the qualitative and quantitative analyses and the implications of these meanings will be discussed in section 8.2.

The knowledge from the analyses of scripts, domain (kernel theories) and justificatory knowledge was used to create the main constructs and models. The iteration process continued to improve the constructs and models development. Using parts of these models, methods and prototype artefacts were developed. Furthermore, the iteration (build and evaluate) process has also been used in developing these two artefacts. The results of the evaluation of these artefacts have shown that the artefacts have performed tasks they were designed for. The knowledge that can be extracted from the development processes of these artefacts and the result of the evaluation and the implication of this knowledge in terms of the research purpose will be discussed in section 8.2.

8.2 Discussing of the findings

Findings 4 and 5 in section 2.7 have discussed the e-assessment inability when assessing subjective questions. One of the main issues for this inability is the absence of students' intention in advance so actions can be taken. The analyses of the scripts showed that even for a human being with the most effort made; it is hard to understand what some of students intended in their workings. This difficulty with other similar challenges has been reported in section 5.2. If a human being (such as the researcher) is facing difficulty in understanding student workings, it is even more difficult for any artefact to perform a task that it is not designed for. Thus, the nature of this topic is a complex problem that requires the approach of complex problem solving as confirmed by both the result of the analyses of the scripts and the literature review.

The discussion of finding 7 in section 2.7 shows that solving a complex problem requires accumulating knowledge about the problem, producing solutions, and having ability to apply the solutions to the problem. This finding was directly related to the need of the analyses of student scripts to produce this knowledge.

The findings in the result of the qualitative analysis contained different types of solution strategies and errors in student workings. The solution strategies are methods that students use to perform a task or tasks. On the other hand, types of errors in the working influence the amount of marks to be awarded to that working.

The findings in the quantitative analysis of scripts added more information to the findings of the qualitative analysis. In the findings of the quantitative analysis, the data about order of operation can be used to inform how often students follow the expected sequence when performing the task of element elimination. For example, in this quantitative analysis, during the element elimination process of a 3 by 3 matrix, the fourth, the seventh and the eighth elements were selected in the expected order by 98.1%, 98.7% and 98.7% of students respectively. The quantitative analysis has also shown that 96.3%, 91.9% and 71.9% of students have used the expected solution strategies when eliminating the fourth, the seventh and the eighth elements respectively.

The findings of the quantitative analysis of student errors have been given in sections 5.5.2 and 5.5.3. These findings are about types of errors, frequencies of their occurrences and places that produced them. In total there were 183 errors, and eight different types of errors that were gathered from six different places in student workings. The details of each type of error, the frequency of its occurrence and the place that produced It, have been given in section 5.5.2, while section 5.5.3 has discussed the meanings these error data.

Having the knowledge of types of student errors would help to recognise these errors in their workings. The knowledge of the places that produce these errors would help identifying the sources of these errors. This might indicate that students lack the domain knowledge or there might be other reasons such as miscopying the original question. For example, for the 'original question miscopy' error the place of its occurrence is at the start of student workings. It is important to check the existence of this error and respond to it. In the analysis of the scripts, the number of students who miscopied the original question was 14%. In other words, out of 160 scripts that have been analysed quantitatively, 22 scripts (each script representing one student) contained the original miscopy error. The frequencies of these errors would enable prioritising their importance. The process of prioritising these errors would help in making decisions whether to include or ignore these errors when preparing assessment criteria that guide the assessment process.

The knowledge that has been deduced from the analyses of scripts assists in creating a design that improves the ability of DSR artefacts assessing subjective questions which are originated from the context of these student scripts. The high percentage of students following the expected order when eliminating elements and using the expected types of solution strategies adds credibility to producing satisfactory designs provided other relevant knowledge and innovative creativity are used.

The main constructs and models have been created from the outcome of the analyses of the scripts. The arguments that have been made and the evidence that has been provided to theorise these models and constructs have been given in section 5.6. The implication of theorising these models and constructs is that the description and the practice that are represented in these models would help improving e-assessment assessing subjective questions and provide solutions regarding assessing subjective question in the context that has been investigated.

As explained in section 2.4.2, the methods that were used by computer aided assessment (CAA) of mathematics do not include checking for types of student errors and types of solution strategies when they assess student workings. The other finding is that previous e-assessment development, especially CAA, do not adequately address research philosophies which are important as explained in chapter three. The role of research philosophies in determining and implementing relevant methodologies have been emphasised in section 3.3. These methodologies have been used to analyse the scripts to produce knowledge. The methodologies that have been used in building and evaluating this research's artefacts have also been selected based on the discussed research philosophies. For example, regarding the stance of ontology, "it is important to clarify the ontological status of these artefacts of interest" (Gregor & Jones, 2007).

There are four models (shown in Figure 6.4, Figure 6.5, Figure 6.6 and Figure 6.7) that represent different activities in student workings. These models also include awarding marks based on the types of student errors or the absence of errors. Generally, these models represent processes and objects. Processes perform tasks while objects are considered as input used by the processes or they are considered as output of these processes. Objects could be the components of other processes and have attributes.

The prototype implemented the model that is shown in Figure 6.8. This model allows gathering student workings. This model is carefully designed to allow students to show all or some of working(s) that otherwise might be in the student scripts. The working that might be in student scripts have been shown in other models (shown in Figure 6.4, Figure 6.5, Figure 6.6 and Figure 6.7). For example, the model in Figure 6.5 shows element elimination and the unknown value computation process. In this model, the elimination process involves using a solution strategy that might be correct and produces a working. This working might contain errors due to the incorrect using of this strategy. Therefore, in this case, the type of error or errors in the working determines the amount of marks that is awarded to this working. It also shows that the output of this working that has a value might be used in the subsequent step. The information in this model must be collected from students to be analysed in the way it has been described in model 6.5.

The prototype artefact implements the model in Figure 6.8. It collects from students all types of working and save them. However, it does not perform the assessment parts that are shown in the models shown in Figure 6.4, Figure 6.5, Figure 6.6 and Figure 6.7. The prototype implementing the functions in model 6.8 has been given in section 6.3.4.

One of the main issues in this research, as explained in many sections of this thesis, is the absence of the knowledge about intentions of students in their next steps. This issue has caused challenges when designing the user interfaces to collect student workings to implement functions that have been shown in the model in Figure 6.8 . The design of these interfaces required caution to avoid disclosing the domain knowledge that students are asked to provide.

Students are reminded to show their intentions of what they are going to perform and how they are going to perform them. For example, to eliminate an element from a matrix, the prototype asks students to select an element to be eliminated from the list of all the elements that form the matrix. Here, the prototype is not giving any hint to the student which element to select. Then the second requirement (reminded by the prototype) is to provide a strategy that student will use to eliminate the selected element. The prototype does not allow students to record any other working which might be recording the usage of the strategy before providing these two (the element and the strategy) information.

When saving student workings, each task, for example, one element elimination task is saved in a separate file. The method of saving student workings in the prototype allows the workings to be retrieved and updated. This action solves one of the issues that has been found in student workings. This issue was when students redo the same working more than once without crossing out the previous working. However, the prototype does not allow that but updates the whole content by overwriting the previous working.

This approach also helps the assessor retrieve parts of the student workings and assess it separately. Since this is assessing student workings in detail, saving each step of student workings separately (1) ensures the simplicity or ease of assessing student workings to award marks and (2) enables uncovering the places that students show weakness or strength in.

The detailed design of the prototype is based on the knowledge mainly produced by the analyses of student scripts and kernel theories as well as justificatory knowledge. The design knowledge that is displayed in the development of the prototype achieves parts of the purpose that models described in section 6.3.2.

This research is aligned with the goal of the DSR improvement quadrant "to create better solutions in the form of more efficient and effective products, processes, services, technologies, or ideas" (Gregor & Hevner, 2013, p. 346). As reported by Thuan, Drechsler, and Antunes (2019) in DSR, the improvement quadrant may have varied output. This includes, as Thuan et al. explain; "proposed constructs, methods, and models as improvements for an existing problem (Haj-Bolouri et al., 2016; Osterwalder, 2004)" (p. 343). Therefore, the focus of this research on producing design knowledge, constructs and model artefacts and partial implementation of model in the prototype achieves the purpose of a DSR project. March and Smith (1995) state that "there is little argument that novel constructs, models, and methods are viable research results, there is less enthusiasm in the information technology literature for novel instantiations" (p. 260). Hence, it is valid to claim the creation of the artefacts that have been produced, in this research, as novel and these artefacts provide purposeful utility to the environment which the analysed student workings belongs to. Furthermore, these artefacts contribute design knowledge to DSR field.

8.3 Research knowledge claims

In chapter seven, the purpose of evaluation, evaluation methods and attributes of the research artefacts have been discussed. The discussion has shown that these artefacts have satisfactorily achieved the first (as referred to in chapter seven) purpose for which they have been evaluated. The artefacts are also expected to contribute knowledge to the fourth (as referred to in chapter seven) purpose. Based on this evaluation result, the knowledge claim to be made will be discussed in this section. This will also discuss other claims to be made.

The knowledge that will be claimed are (1) extracting solution strategies from scripts, (2) constructs, (3) models, (4) the prototype, (5) the questionnaire. Regarding the analysis of scripts for student errors, there are previous publications that have studied student errors in working. However, to the knowledge of the researcher, there is no previous study that has analysed solution strategies in student workings. Therefore, the act of extracting this type of knowledge as well as the knowledge itself are novel. This knowledge can be utilised in the domain for teaching and assessing and for developing artefacts that might be able to assess subjective questions.

The solution strategies and errors subcategories, in the analyses of the scripts, are constructs. Some errors constructs can be claimed as novel concepts and may be used in improving the e-assessment of mathematics in multi-steps solutions. As shown in Table 5.3, the two errors which are ‘purposeful error’ and ‘ignored operation error’ are constructs which are novel. For these types of errors, multiple usages have been described by Walker et al. (2015) as follow:

> The educational benefits to being able to efficiently capture and report specific errors to students and staff are threefold
>
> - this information may be used to improve the questions by providing enhanced, tailored feedback which will benefit future students taking the e-Assessment;
> - through the use of the re-marking facility current students, who have already tried the e- Assessment, may access this new improved feedback by viewing details of their previous attempts;
> - by looking at the results for a particular cohort, the academic is able to see which areas of the syllabus need more emphasis in lectures. (p. 3)

Greenhow's (2015) view of the usage of these types of errors in designing e-assessment for objective assessment has been stated at the start of section 8.1.

The models as the main artefact besides constructs can also be claimed as the purposeful and novel artefact. The usefulness of these models in describing the context that has been studied and providing solutions has been explained in sections 5.2, 5.3 and 5.4. Moreover, the discussion about theorising these models has been given in section 5.6. The models reaching the level of theory, adds weight to the credibility of claiming the models as significant contribution to the environment that produced the student workings. The methods of qualitative analysis and of model design are generic, so the same methods could be equally applied to any solutions of multi-steps questions in any discipline. The knowledge in these methods are transferable and can achieve some of the fourth purpose of evaluation that has been mentioned at the start of this section. The prototype is also qualified to be claimed as the useful artefact implementing one of these models.

The other important output of this research is the questionnaire instrument that was created to perform the quantitative analysis of the scripts. The role of this questionnaire development process in improving the findings in the qualitative analysis of scripts has been explained in different sections of this thesis, especially in section 4.5.3. The absence of the modifiable questionnaire has led to the creation of this new one. This questionnaire could be used in a similar field to analyse solution strategies that involve questions with multi-steps solutions. In this research, the solution strategies that have been uncovered during the qualitative analysis of scripts have been categorised and given names which are not specific to the discipline. For example, if a solution strategy in a student working is 2R2 + R3 and it is an ‘expected correct’ solution strategy for the task, then this solution strategy in the student workings is categorised under the ‘expected’ solution category. The quantitative analysis of scripts collected the quantified data about solution strategies and analysed it. This quantified data about solution strategies is the qualitative data in the form of ‘expected’, ‘acceptable’ and ‘incorrect’ categories. Hence, this approach of categorisation is general and with little modification the questionnaire can be used in a related research field.

In summary, the knowledge base and the researchers’ judgement are the sources that assist in (1) revealing sections or activities that contain the contributed knowledge, (2) forming the knowledge validation criteria and (3) selecting appropriate knowledge acquisition methods or creating new methods to extract knowledge.

The knowledge from the student scripts has been produced using inductive, deductive, abstraction and reflection methods. The contributed design science knowledge in the DSR output is obtained using reflection and abstraction cognitions Vaishnavi et al. (2017). Abstraction cognition process has been covered in section 4.4.4. The reflection cognition process is explained next.

Reflection is important as a process for extracting and evaluating knowledge. Redmond (2014) describes the action of reflection being dependent on using various types of thinking and those types have given by Lipman (2003) as “recursive thinking, metacognitive thinking, self-corrective thinking" (p. 27). According to Lundgren et al. (2017), the purpose of reflection is to achieve meaning by employing mindful and specific process that entails the initiation of reflective activities "such as noticing, reviewing, interpreting, framing, generalizing, and applying metacognition" (p. 307). When implementing the reflective activities, the reflection process requires observers to apply measured and critical analysis in the interpretive view they make and in the philosophical assumptions they take (Erickson, 1985). Observers' application of credible analysis demands the employment of cognisant reflection on the entire research activities to identify the research contributed knowledge; these activities include tasks such as problem formulation, appropriate methodology selection and the interpretation of the research efforts outcome (Sein, Henfridsson, Purao, Rossi, & Lindgren, 2011). Hence, the application of the reflection process, in research, is a mean of acquiring knowledge that researchers use to enable them to make and to relate valuable meaning and knowledge. More details of knowledge acquisition methods have been explained in sections 3.3.3, 4.2 and 4.4.4.

Researchers as being one of the research instruments are required to "make choices regarding data sources, data construction, and analysis methods that best fit their research questions and to consider using multiple approaches and modes of inquiry" (Ercikan & Roth, 2006, p. 55). This shows researchers influence and impacts on the whole processes of research and beyond (Salmona & Smith, 2014). Hence, to claim the contribution of the knowledge that has been discussed in this section, the researcher has used appropriate knowledge acquisition research philosophies, methodologies and methods.

To validate the knowledge contributed, representational fidelity and implementability criteria (Hevner et al., 2004) have been used. These two criteria have been demonstrated in different sections of this thesis.

8.4 Limitations in this research

One of this research main purposes was to understand the difficulty of assessing subjective assessment using e-assessment of mathematics. Since this is a wide and complex field, the review focused on the knowledge produced by subjective assessment, ways of assessing subjective questions by e-assessment tools and methodologies of developing e-assessment. There are too many existing systems of this kind to examine them. This shows the limitation to claim with confidence the method of their development and functionalities they support. However, the review was able to uncover their general development and their inability to assess subjective questions and this was one of the main purposes of the review.

The criteria for selecting scripts meant there was a limited number available for analysis. The criteria to include the student scripts in the analyses was based on the student workings showing 20% or above of the total marks that was assigned to that particular solution. These criteria reduced the number of scripts to be analysed. Consequently, this reduces the probability of uncovering more types of solution strategies and errors from student workings.

The selected scripts were initially analysed using the inductive qualitative analysis. The problem with this approach, as some researchers claim, involves researchers' bias in the analysis process. The criticism of an inductive approach is that researchers who use this approach tend to use their prior knowledge when they investigate context or subject (Chapman, Hadfield, & Chapman, 2015) and "each researcher brings a personal bias into their individual perceptions of what they are observing" (Soiferman, 2010, p. 13).

However, in the analysis of student workings, there was no better approach, at this stage, to analyse scripts. At the initial stage, when the scripts were analysed, the main purpose was to find out what could be extracted from the student workings without involving interpretation. Initially a thorough investigation of research philosophies including axiology was made. The understanding of the roles of subjectivism and objectivism in conducting research helped the selection of the most appropriate methodologies and methods. The researcher's own understanding and belief in the effects of value-laden and value-free concepts on the research process and outcomes has also minimised the opportunity for bias. The evaluation and validation of the qualitative analysis process and the outcome of the process, explained in section 5.1, have further reduced the effect of this bias. Additionally, the process of creating the questionnaire improved the outcome of the qualitative analysis. The difficulty in understanding some student workings during the analysis process and the challenge of categorising some findings also affected the outcome of the analysis. This issue was explained in section 5.2.

The other limitation was related to the type of the quantitative analysis of the scripts. In this phase, the only frequency analysis was performed. The result of this analysis has added important information to the findings in the qualitative analysis, and further statistical analysis would not alter or affect the type of solution strategies and errors that have been created from the findings of the qualitative phase.

The limitations that have been associated with the analyses of the scripts have also affected the constructs and models artefacts, since these artefacts have been built using the knowledge that has been deduced from the analyses of the scripts.

The evaluation methods that were used to evaluate the prototype artefacts might also have limitations. The debugging evaluation method might not be able to evaluate the performance and the efficiency of the prototype. However, the observational evaluation method that was used with the testing method might add more value to the result of the evaluation but might not be sufficient to claim that a complete evaluation has been performed.

The models artefact although they are generic and theorised, might also lack the completeness as explained in chapter seven when discussed the attributes of this DSR project's artefacts.

8.5 Further directions

Future work of this research may discuss the contributed knowledge. Solution strategies can be analysed in different disciplines that use questions with multi-steps solutions to verify the transferability of the method of categorisation and naming as used in this research. The discussion here is not about the approach of the analysis which used a qualitative contents approach. It is about ways of categorising the findings of types of solution strategies using terms such as ‘expected’, ‘acceptable’ and ‘incorrect’. These categorised (uncovered) strategies might be stored, for future use, in a repository that is accessible by others.

The models that have been theorised can be tested using the deductive approach to verify the claimed models as theories. Furthermore, the models can be generalised, so that they can be applied to solutions in different topics or discipline.

The questionnaire can also be tested or modified and used in other contexts.

The other main issue is the user interface in the prototype. This interface can be improved in terms of hiding the knowledge that student should show for awarding marks. At the same time the interface should be able to guide students to perform the various tasks which constitute the solution. This may also be divided into sections, and this division helps students to show (save) their solution strategies and usage of these strategies separately. This approach of collecting student solution of a problem should also assist assessors to retrieve these solutions and award marks based on the correctness of solution strategies, details of using the strategies and types of errors that might influence the amount of marks.

8.6 Conclusion

The importance of authentic assessment in assisting students to acquire lasting knowledge, and which contributes to the success of their professional lives, has been established. This lasting effect is the result of many factors among which subjective assessment is one. The inability of e-assessment to perform subjective assessment provides an opportunity for investigation.

This research performed rigorous analyses of student's workings that contain solutions of multi-steps questions as well as conducting thorough literature reviews to understand the complexity of assessing subjective questions using e-assessment for mathematics systems. These two processes produced knowledge that could be used in multiple disciplines including the environment which was the source of all student's workings. In addition, this knowledge includes design knowledge that could significantly improve the development of DSR artefacts that will be used in assessing subjective questions.

References

Adesina, A., Stone, R., Batmaz, F., & Jones, I. (2014). Touch Arithmetic: A process-based Computer- Aided Assessment approach for capture of problem solving steps in the context of elementary mathematics. *Computers & Education, 78*, 333-343.

Adesina, A. O., Stone, R. G., Batmaz, F., & Jones, I. (2015). A semi-automatic computer-aided assessment approach for marking and providing feedback comments.

adore Newman, I., Ridenour, C. S., Newman, C., & DeMarco Jr, G. M. P. (2003). A apology of research purposes and its relationship to mixed methods. *Handbook of mixed methods in social & behavioral research*, 167.

Afamasaga-Fuata'i, K. (2008). Students' conceptual understanding and critical thinking: A case for concept maps and vee-diagrams in mathematics problem solving. *Australian Mathematics Teacher, The, 64*(2), 8.

Ågerfalk, P. J. (2013). Embracing diversity through mixed methods research: Taylor & Francis. Aier, S., & Fischer, C. (2011). Criteria of progress for information systems design theories. *Information Systems and E-Business Management, 9*(1), 133-172.

Aken, J. E. v. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies, 41*(2), 219-246.

Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. *European Journal of Information Systems, 19*(3), 359-376.

Aldawod, A., & Day, J. (2017). A Critical Reflection upon the Postmodernist Philosophical Positions and Issues Relevant to Entrepreneurship Research.

Alexander, N. (2015). What’s more general than a whole population? *Emerging Themes in Epidemiology, 12*, 11. doi:10.1186/s12982-015-0029-4

Algebra - Augmented Matrices - Pauls Online Math Notes. Section 7-3 : Augmented Matrices. Retrieved from http://tutorial.math.lamar.edu/Classes/Alg/AugmentedMatrix.aspx

Alter, S. (2008). Defining information systems as work systems: implications for the IS field. *European Journal of Information Systems, 17*(5), 448-469.

Alturki, A., Bandara, W., & Gable, G. G. (2012). *Design science research and the core of information systems.* Paper presented at the International Conference on Design Science Research in Information Systems.

Alturki, A., Gable, G. G., & Bandara, W. (2013). *BWW ontology as a lens on IS design theory: extending the design science research roadmap.* Paper presented at the International Conference on Design Science Research in Information Systems.

Arifin, S., Putri, R., Hartono, Y., & Susanti, E. (2017). *Developing Ill-defined problem-solving for the context of “South Sumatera”.* Paper presented at the Journal of Physics: Conference Series.

Arnott, D., & Pervan, G. (2012). Design science in decision support systems research: An assessment using the Hevner, March, Park, and Ram Guidelines. *Journal of the Association for Information Systems, 13*(11), 923-949.

Ashton, H. S., Beevers, C. E., Korabinski, A. A., & Youngson, M. A. (2006). Incorporating partial credit in computer-aided assessment of Mathematics in secondary education. *British Journal of Educational Technology, 37*(1), 93-119. doi:10.1111/j.1467-8535.2005.00512.x

Asikainen, H. (2014). Successful learning and studying in Biosciences: Exploring how students conceptions of learning, approaches to learning, motivation and their experiences of the teaching-learning environment are related to study success.

Asplin, G. (2018). *Traffic Light System-BasicADL (TLS-BasicADL) Development, reliability, validity, clinical utility and patient perspective*.

Astroth, K. S. (2018). Exploring the Evidence Quantative and Qualitative Research: Focusing on the Fundamentals: Reading Qualitative Research with a Critical Eye. *Nephrology Nursing Journal, 45*(4), 381-386.

Balcombe, A., Brennan, M., & Everiss, L. (2011). *Formative Learning: A Description of an Assessment Framework in Initial Teacher Education.* Paper presented at the Proceedings of the Symposium on Assessment and Learner Outcomes, Victoria University.

Bansal, P., Smith, W. K., & Vaara, E. (2018). New Ways of Seeing through Qualitative Research. *Academy of Management Journal, 61*(4), 1189-1195. doi:10.5465/amj.2018.4004 Bartsch, H., Thompson, W., Jernigan, T., & Dale, A. (2014). A web-portal for interactive data exploration, visualization, and hypothesis testing. *Frontiers in Neuroinformatics, 8*(25). doi:10.3389/fninf.2014.00025

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artefact and Theory. *Journal of the Association for Information Systems, 19*(5), 358-376.

Bayazit, N. (2004). Investigating design: A review of forty years of design research. *Design issues, 20*(1), 16-29.

Bazeley, P. (2004). Issues in mixing qualitative and quantitative approaches to research. *Applying qualitative methods to marketing management research, 141*, 156.

Beckman, S., Nilsson, G. B., & Dahlbom, B. (2002). Artefacts and artificial science: Almqvist & Wiksell International.

Beese, J., Haki, K., & Aier, S. (2015). On the Conceptualization of Information Systems as Socio- Technical Phenomena in Simulation-Based Research.

Beevers, C., Wild, D., McGuine, G., Fiddes, D., & Youngson, M. (1999). Issues of partial credit in mathematical assessment by computer. *ALT-J, 7*(1), 26-32.

Bellamy, A. (2015). A critical analysis of teaching and learning negative numbers. *Philosophy of Mathematics Education Journal, 29*.

Belli, G. (2009). *Nonexperimental quantitative research*: na.

Bernal, D. D. (2002). Critical race theory, Latino critical theory, and critical raced-gendered epistemologies: Recognizing students of color as holders and creators of knowledge. *Qualitative inquiry, 8*(1), 105-126.

Biedenbach, T., & Jacobsson, M. (2016). The open secret of values: the roles of values and axiology in project research. *Project Management Journal, 47*(3), 139-155.

Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research. *Sage handbook of mixed methods in social and behavioral research, 2*, 95-118.

Bischof, N., & Eppler, M. J. (2011). Caring for Clarity in Knowledge Communication. *J. UCS, 17*(10), 1455-1473.

Black, P., & Wiliam, D. (2005). Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal, 16*(2), 249-261. doi:10.1080/09585170500136218

Blaikie, N. (2007). *Approaches to social enquiry: Advancing knowledge*: Polity.

Blair, K. P., Rosenberg-Lee, M., Tsang, J. M., Schwartz, D. L., & Menon, V. (2012). Beyond natural numbers: negative number representation in parietal cortex. *Frontiers in human neuroscience, 6*, 7.

Blessing, L. T., & Chakrabarti, A. (2009). *DRM: A Design Reseach Methodology*: Springer.

Bloxham, S., & Boyd, P. (2007) *Developing Effective Assessment in Higher Education A Practical Guide* (1 ed., pp. 17-17): McGraw-Hill Education (UK).

Boekaerts, M., Seegers, G., & Vermeer, H. (1995). Solving math problems: Where and why does the solution process go astray? *Educational Studies in Mathematics, 28*(3), 241-262.

Booth, J. L., Barbieri, C., Eyer, F., & Paré-Blagoev, E. J. (2014). Persistent and pernicious errors in algebraic problem solving. *The Journal of Problem Solving, 7*(1), 3.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal, 9*(2), 27-40.

Bracken, S. (2010). Discussing the Importance of Ontology and Epistemology Awareness in Practitioner Research. *Worcester Journal of Learning and teaching*(4), 1-9.

Bryman, A. (2008). *Social research methods*: Oxford ; New York : Oxford University Press 3rd ed.

Bunge, M. (1979). "Philosophical Inputs and Outputs of Technology". In R. C. Scharff & a. V. Dusek (Eds.), *Philosophy of Technology* (2 ed., pp. 191-200): John Wiley & Sons, Incorporated.

Burns, N., & Grove, S. K. (2010). *Understanding nursing research: Building an evidence-based practice*: Elsevier Health Sciences.

Cambridge Dictionary online. Retrieved from https://dictionary.cambridge.org/dictionary/english/mark

Cangelosi, R., Madrid, S., Cooper, S., Olson, J., & Hartter, B. (2013). The negative sign and exponential expressions: Unveiling students' persistent errors and misconceptions. *The Journal of Mathematical Behavior, 32*(1), 69-82.

Chapman, A., Hadfield, M., & Chapman, C. (2015). Qualitative research in healthcare: an introduction to grounded theory using thematic analysis. *Journal of the Royal College of Physicians of Edinburgh, 45*(3), 201-205.

Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer.

Chen, Y.-Y., Shek, D. T., & Bu, F.-F. (2011). Applications of interpretive and constructionist research methods in adolescent research: philosophy, principles and examples. *International journal of adolescent medicine and health, 23*(2), 129-139.

Chilivumbo, C. (2015). Question Banks: A tool for improving Higher Education Assessment across National Resource Networks: The Polytechnic of Malawi case study.

Clark, A. M., MacIntyre, P. D., & Cruickshank, J. (2007). A critical realist approach to understanding and evaluating heart health programmes. *Health:, 11*(4), 513-539.

Collopy, F., & Boland, R. (2004). *Managing as designing*: Stanford University Press Redwood City, CA. Conole, G., & Warburton, B. (2005). A review of computer-assisted assessment. *ALT-J, 13*(1), 17-31.

Cook, J., & Jenkins, V. (2010). Getting started with e-assessment.

Costa, E., Soares, A. L., & de Sousa, J. P. (2016). *Situating case studies within the design science research paradigm: An instantiation for collaborative networks.* Paper presented at the Working Conference on Virtual Enterprises.

Creswell, J. (2003). *Research Design*.

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative* (4th ed.): Pearson.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications.

Creswell, J. W. (2014). *Research design : qualitative, quantitative, and mixed methods approaches*: Thousand Oaks, California : SAGE Publications, Inc., 2014.

Fourth edition, international student edition.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*: Sage publications.

Creswell, J. W., & Creswell, J. W. (2013). *Qualitative inquiry & research design : choosing among five approaches*: Thousand Oaks : SAGE Publications

3rd ed.

Cross, N. (2006). *Designerly ways of knowing*: London : Springer, 2006.

Crotty, M. (1998). The foundations of social research: meaning and Perspective in the Research Process.

Dahlbom, B., Beckman, S., & Nilsson, G. B. (2002). *Artefacts and Artificial Science*.

De Smith, M. (2015). STATSREF: Statistical Analysis Handbook-a web-based statistics.".

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*: Thousand Oaks : Sage, [2011] 4th ed.

Dhokrat, A., & Mahender, C. N. (2012). Automated Answering for Subjective Examination. *International Journal of Computer Applications, 56*(14).

Diagonal Matrices Upper and Lower Triangular Matrices - Faculty. Diagonal Matrices, Upper and Lower Triangular Matrices Linear Algebra. Retrieved from http://faculty.etsu.edu/joynerm/Diagonal%20Matrices.pdf Dörner, D., & Funke, J. (2017). Complex problem solving: what it is and what it is not. *Frontiers in psychology, 8*, 1153.

Drechsler, A., & Hevner, A. (2016). *A four-cycle model of IS design science research: capturing the dynamic nature of IS artefact design.* Paper presented at the Breakthroughs and Emerging Insights from Ongoing Design Science Projects: Research-in-progress papers and poster presentations from the 11th International Conference on Design Science Research in Information Systems and Technology (DESRIST) 2016. St. John, Canada, 23-25 May.

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological monographs, 58*(5), i.

Ebrahimi Belil, F., Alhani, F., Ebadi, A., & Kazemnejad, A. (2018). Self-efficacy of people with chronic conditions: A qualitative directed content analysis. *Journal of clinical medicine, 7*(11), 411.

El Idrissi, S. C., & Corbett, J. (2016). Green IS research: A modernity perspective. *CAIS, 38*, 30.

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative Content Analysis. *SAGE Open, 4*(1), 2158244014522633. doi:doi:10.1177/2158244014522633

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing, 62*(1), 107-115.

Ercikan, K., & Roth, W.-M. (2006). What good is polarizing research into qualitative and quantitative?

Educational Researcher, 35(5), 14-23.

Erickson, F. (1985). *Qualitative methods in research on teaching*: Citeseer.

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine, 7*(3), 93-99.

Evely, A. C., Fazey, I. R. A., Pinard, M., & Lambin, X. (2008). The influence of philosophical perspectives in integrative research: a conservation case study in the Cairngorms National Park. *Ecology and Society*.

Faithfull, W. J., Rodríguez, J. J., & Kuncheva, L. I. (2019). Combining univariate approaches for ensemble change detection in multivariate data. *Information Fusion, 45*, 202-214.

Faste, T., & Faste, H. (2012). *Demystifying "design research": Design is not research, research is design.* Paper presented at the IDSA Education Symposium.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods, 5*(1), 80-92.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health services research, 48*(6pt2), 2134-2156.

Fine, A. H. (2010). *Handbook on Animal-Assisted Therapy : Theoretical Foundations and Guidelines for Practice*. Saint Louis, UNITED STATES: Elsevier Science & Technology.

Fischer, A., Greiff, S., & Funke, J. (2012). The Process of Solving Complex Problems. *The Journal of Problem Solving, 4*(1). doi:10.7771/1932-6246.1118

Forman, J., & Damschroder, L. (2008). Qualitative content analysis. *Empirical Research for Bioethics: A Primer. Oxford, UK: Elsevier Publishing*, 39-62.

Fuadiah, N. F., & Suryadi, D. (2017). Some Difficulties in Understanding Negative Numbers Faced by Students: A Qualitative Study Applied at Secondary Schools in Indonesia. *International Education Studies, 10*(1), 24-38.

Funke, J. (2010). Complex problem solving: a case for complex cognition? *Cogn Process, 11*(2), 133- 142. doi:10.1007/s10339-009-0345-0

Gallardo, A. (1995). Negative Numbers in the Teaching of Arithmetic. Repercussions in Elementary Algebra.

Gears, C. (2005). Classroom Assessment: Minute by Minute, Day by Day. *Assessment, 63*(3).

Gelso, C. J. (2006). Applying Theories to Research. *The psychology research handbook: A guide for graduate students and research assistants*, 455.

Gettelman, A., & Rood, R. B. (2016). Model Evaluation *Demystifying Climate Models: A Users Guide to Earth System Models* (pp. 161-176). Berlin, Heidelberg: Springer Berlin Heidelberg.

Gibbs, G. (1999). Using Assessment Strategically to Change the Way Students. *Assessment matters in higher education*, 41.

Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference? *Educational studies, 32*(4), 399-409.

Gorard, S. (2010). Research design, as independent of methods.

Gorelick, R. (2011). What is theory? *Ideas in Ecology and Evolution, 4*.

Gray, D. E. (2004). Theoretical perspectives and research methodologies. *Gray, DE Doing*.

Greene, J. C., & Hall, J. N. (2010). Dialectics and pragmatism: Being of consequence. *Sage handbook of mixed methods in social & behavioral research*, 119-143.

Greenhow, M. (2015). Effective computer-aided assessment of mathematics; principles, practice and results. *Teaching Mathematics and Its Applications: International Journal of the IMA, 34*(3), 117-137.

Gregg, D. G., Kulkarni, U. R., & Vinzé, A. S. (2001). Understanding the philosophical underpinnings of software engineering research in information systems. *Information Systems Frontiers, 3*(2), 169-183.

Gregor, S. (2010). Building theory in a practical science. *Information systems foundations: the role of design science (Eds Gregor, S. and Hart, D.). Australian National University Press, Canberra, Australia*, 51-74.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly, 37*(2), 337-356.

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information systems, 8*(5), 312-335.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Ectj, 29*(2), 75. Guba, E. G. (1990). *The Paradigm dialog*: Newbury Park, Calif. : Sage Publications.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research, 2*(163-194), 105.

Guest, G., Namey, E. E., & Mitchell, M. L. (2012). *Collecting Qualitative Data: A Field Manual for Applied Research*: SAGE Publications.

Guest, G., Namey, E. E., & Mitchell, M. L. (2013). *Collecting qualitative data : a field manual for applied research*: Thousand Oaks : SAGE Publications.

Gyll, S., & Ragland, S. (2018). Improving the validity of objective assessment in higher education: Steps for building a best-in-class competency-based assessment program. *The Journal of Competency-Based Education, 3*(1), e01058.

Hanson, J. (2011). *A determination and classification of student errors in lower-level calculus through computer-aided assessment and analysis.* Brunel University, School of Information Systems, Computing and Mathematics.

Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. *Journal of counseling psychology, 52*(2), 224.

Hardin, L. E. (2003). Problem-solving concepts and theories. *Journal of veterinary medical education, 30*(3), 226-229.

Hathaway, R. S. (1995). Assumptions underlying quantitative and qualitative research: Implications for institutional research. *Research in higher education, 36*(5), 535-562.

Herselman, M., & Botha, A. (2015). *Evaluating an artefact in design science research.* Paper presented at the Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists.

Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol.22): Springer Science & Business Media.

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian journal of information systems, 19*(2).

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly, 28*(1), 75-105.

Holden, M. T., & Lynch, P. (2004). Choosing the appropriate methodology: understanding research philosophy. *The marketing review, 4*(4), 397-409.

Holmberg, C., Larsson, C., Korp, P., Lindgren, E.-C., Jonsson, L., Fröberg, A., . . . Berg, C. (2018). Empowering aspects for healthy food and physical activity habits: adolescents' experiences of a school-based intervention in a disadvantaged urban community. *International journal of qualitative studies on health and well-being, 13*(sup1), 1487759.

Hooks, B. (2000). *Feminist theory: From margin to center*: Pluto Press.

Hu, X., & Xia, H. (2010). *Automated assessment system for subjective questions based on LSI.* Paper presented at the Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on.

Huba, M. E., & Freed, J. E. (2000). Learner centered assessment on college campuses: Shifting the focus from teaching to learning. *Community College Journal of Research and Practice, 24*(9), 759-766.

Hudson, L. A., & Ozanne, J. L. (1988). Alternative ways of seeking knowledge in consumer research. *Journal of consumer research, 14*(4), 508-521.

Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian journal of information systems, 19*(2), 5.

Jain, S. K., & Gunawardena, A. (2004). *Linear algebra: an interactive approach*: Thompson- Brooks/Cole.

Jannaber, S., Riehle, D. M., Delfmann, P., Thomas, O., & Becker, J. (2017). *Designing a Framework for the Development of Domain-Specific Process Modelling Languages.* Paper presented at the International Conference on Design Science Research in Information Systems.

Jeusfeld, M. A., Jarke, M., Nissen, H. W., & Staudt, M. (1998). ConceptBase: Managing conceptual models about information systems *Handbook on Architectures of Information Systems* (pp. 273-294): Springer.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher, 33*(7), 14-26.

Jordan, S. (2014). *E-assessment for learning? Exploring the potential of computer-marked assessment and computer-generated feedback, from short-answer questions to assessment analytics.* The Open University.

Keeling, R. P., & Dungy, G. J. (2004). *Learning reconsidered : a campus-wide focus on the student experience*: Washington, D.C : ACPA : NASPA, c2004.

Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*: Unesco, International Institute for Educational Planning.

Ken, P., Tuure, T., Marcus, A. R., & Samir, C. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*(3), 45. doi:10.2753/MIS0742-1222240302

Ketterlin-Geller, L. R., & Yovanoff, P. (2009). Diagnostic assessments in mathematics to support instructional decision making. *Practical Assessment, Research & Evaluation, 14*(16), 1-11.

Kilhamn, C. (2011). *Making sense of negative numbers*: Department of Pedagogical, Curricular and Professional Studies ….

Kirkwood, B. R., & Sterne, J. A. (2010). *Essential medical statistics*: John Wiley & Sons.

Kivunja, C., & Kuyini, A. B. (2017). Understanding and Applying Research Paradigms in Educational Contexts. *International Journal of Higher Education, 6*(5), 26-41.

Korpela, M., Mursu, A., Soriyan, A., Eerola, A., Häkkinen, H., & Toivanen, M. (2004). Information systems research and development by activity analysis and development: dead horse or the next wave? *Information systems research* (pp. 453-471): Springer.

Kothari, C. R. (2004). *Research methodology: Methods and techniques*: New Age International.

Kuechler, W., & Vaishnavi, V. (2012). A framework for theory development in design science research: multiple perspectives. *Journal of the Association for Information Systems, 13*(6), 395.

Kuhn, T. S. (1970). *The structure of scientific revolutions*: University of Chicago press.

Kulasegaram, K., & Rangachari, P. K. (2018). Beyond “formative”: assessments to enrich student learning. *Advances in physiology education, 42*(1), 5-14.

Kullberg, A. (2010). *What is taught and what is learned. Professional insights gained and shared by teachers of mathematics*: Department of Pedagogical, Curricular and Professional Studies ….

Kumaran, V. S., & Sankar, A. (2015). Towards an automated system for short-answer assessment using ontology mapping. *International Arab Journal of e-Technology, 4*(1), 17-24.

Laura M. O'Dwyer, & James A. Bernauer. (2014). Non-Experimental Research Designs *Quantitative Research for the Qualitative Researcher* (pp. 152-170). Thousand Oaks: SAGE Publications, Inc.

Lawson, D. (2012). Computer-aided assessment in mathematics: Panacea or propaganda?

International Journal of Innovation in Science and Mathematics Education (formerly CAL- laborate International), 9(1).

Lee, A. S. (2014). Theory is king? But first, what is theory? *Journal of information technology, 29*(4), 350-352. doi:10.1057/jit.2014.23

Leedy, P. D., & Ormrod, J. E. (2012). *Practical research : planning and design*: Upper Saddle River, N.J. ; Harlow : Pearson Education 10th ed., International ed.

Lehman, A., O'Rourke, N., Hatcher, L., & Stepanski, E. (2013). *JMP for basic univariate and multivariate statistics: Methods for researchers and social scientists*: SAS Institute.

Lilley, M., Barker, T., & Maia, M. (2002). *THE USE OF OBJECTIVE QUESTIONS IN HIGHER EDUCATION: POTENTIAL AND LIMITATIONS*.

Lipman, M. (2003). *Thinking in education*: Cambridge University Press.

Liu, Z., & Jansen, B. J. (2015). *Subjective versus Objective Questions: Perception of Question Subjectivity in Social Q&A*, Cham.

Livne, N. L., Livne, O. E., & Wight, C. A. (2007). Can automated scoring surpass hand grading of students' constructed responses and error patterns in mathematics. *MERLOT Journal of Online Learning and Teaching, 3*(3), 295-306.

Luckin, R., Clark, W., Avramides, K., Hunter, J., & Oliver, M. (2017). Using teacher inquiry to support technology-enhanced formative assessment: a review of the literature to inform a new method. *Interactive Learning Environments, 25*(1), 85-97.

Lundgren, H., Bang, A., Justice, S. B., Marsick, V. J., Poell, R. F., Yorks, L., . . . Sung, S. (2017). Conceptualizing reflection in experience-based workplace learning. *Human Resource Development International, 20*(4), 305-326.

Mabuza, L. H., Govender, I., Ogunbanjo, G. A., & Mash, B. (2014). African Primary Care Research: Qualitative data analysis and writing results. *African journal of primary health care & family medicine, 6*(1), 1-5.

Mack, L. (2010). The philosophical underpinnings of educational research: Polyglossia.

Madsen, B. (2011). *Statistics for non-statisticians*: Berlin ; Heidelberg ; New York : Springer.

MAPLESOFT. (2016). Challenges and Solutions in Automated STEM Assessment. Retrieved from https://www.maplesoft.com/products/mapleta/

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision support systems, 15*(4), 251-266.

Marwedel, P. (2011). Evaluation and Validation *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems* (pp. 203-234). Dordrecht: Springer Netherlands.

Maxwell, J. (2012). What is realism, and why should qualitative researchers care. *A realist approach for qualitative research*, 3-13.

Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. *Handbook of mixed methods in social & behavioral research*, 145-168.

Mayring, P. (2000). Qualitative Content Analysis. *2000, 1*(2).

Mayring, P. (2014). Qualitative content analysis: theoretical foundation, basic procedures and software solution.

McCarthy, J. (1980). Circumscription—a form of non-monotonic reasoning. *Artificial intelligence, 13*(1-2), 27-39.

McCarthy Veach, P., LeRoy, B. S., & Callanan, N. P. (2018). Gathering Information: Asking Questions *Facilitating the Genetic Counseling Process: Practice-Based Skills* (pp. 111-137). Cham: Springer International Publishing.

McClarty, K. L., & Gaertner, M. N. (2015). Measuring Mastery: Best Practices for Assessment in Competency-Based Education. AEI Series on Competency-Based Higher Education. *American Enterprise Institute for Public Policy Research*.

McKenney, J. L., & Keen, P. G. (1974). How managers' minds work. *Harvard Business Review, 52*(3), 79-90.

Messerschmidt, J. A. (2016). Foundation of Academic Assessment.

Mingers, J. (2001). Combining IS research methods: towards a pluralist methodology. *Information systems research, 12*(3), 240-259.

Mink, O. G. (1992). Creating new organizational paradigms for change. *International Journal of Quality & Reliability Management, 9*(3).

Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. *Conservation Biology, 28*(5), 1167-1177.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1), 48-76.

Morgan, J. (2014). Critical realism as a social ontology for economics. *Handbook of Research Methods and Applications in Heterodox Economics*.

Mouza, M. (2018). *WHAT IS THEORY*.

Muratovski, G. (2015). Paradigm shift: Report on the new role of design in business and society. *She Ji: The Journal of Design, Economics, and Innovation, 1*(2), 118-139.

Mwilu, O. S., Comyn-Wattiau, I., & Prat, N. (2016). Design science research contribution to business intelligence in the cloud—A systematic literature review. *Future Generation Computer Systems, 63*, 108-122.

Newman, I., Ridenour, C., Newman, C., & De Marco, G. (2002). A typology of research purposes and its relationship to mixed methods in social and behavioral research (pp. 167-188). *Thousand Oaks, CA, EE. UU.: Sage*.

Nicolas, H., & Liora, L. (1994). A Cognitive Gap between Arithmetic and Algebra. *Educational Studies in Mathematics, 27*(1), 59.

Niehaves, B. (2007). On episemological diversity in design science: New vistas for a design-oriented IS research? *ICIS 2007 Proceedings*, 133.

Nunamaker, J. F., Chen, M., & Purdin, T. D. M. (1990). Systems Development in Information Systems Research. *Journal of Management Information Systems, 7*(3), 89-106. doi:10.1080/07421222.1990.11517898

Offermann, P., Blom, S., Levina, O., & Bub, U. (2010). Proposal for Components of Method Design Theories. *Business & Information Systems Engineering, 2*(5), 295-304. doi:10.1007/s12599- 010-0120-x

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information systems research, 2*(1), 1-28.

Osterwalder, A. (2004). The business model ontology: A proposition in a design science approach.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(5), 533-544.

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research & Evaluation, 13*(4), 1-11.

Papajorgji, P. J., & Pardalos, P. M. (2014). Use Cases and Actors *Software Engineering Techniques Applied to Agricultural Systems* (pp. 47-52): Springer.

Patton, M. Q. (2002). *Qualitative research & evaluation methods / by Michael Quinn Patton*. Thousand Oaks [u.a.]: Sage.

Payne, S. J., & Squibb, H. R. (1990). Algebra mal-rules and cognitive accounts of error. *Cognitive science, 14*(3), 445-481.

Pead, D. (2010). *On computer-based assessment of mathematics.* University of Nottingham. Peffers, K., Rothenberger, M., Tuunanen, T., & Vaezi, R. (2012). *Design Science Research Evaluation*, Berlin, Heidelberg.

Pereira, R., Almeida, R., & da Silva, M. (2013). How to Generalize an Information Technology Case Study. In J. vom Brocke, R. Hekkala, S. Ram, & M. Rossi (Eds.), *Design Science at the Intersection of Physical and Virtual Design* (Vol. 7939, pp. 150-164): Springer Berlin Heidelberg.

Pereira, R., Almeida, R., & da Silva, M. M. (2013). How to generalize an information technology case study *Design Science at the Intersection of Physical and Virtual Design* (pp. 150-164): Springer.

Perfect, C. (2015). *A demonstration of Numbas, an e-assessment system for mathematical disciplines.* Paper presented at the CAA Conference.

Pham, H. (2016). Security demands, organisational and personal resources: A stress-based security compliance model.

Pierce, J. (2014). *On the presentation and production of design research artefacts in HCI.* Paper presented at the Proceedings of the 2014 conference on Designing interactive systems.

Porta, M. (2014). *A dictionary of epidemiology*: Oxford university press.

Prasad, B. D. (2008). Content analysis. *Research methods for social work, 5*, 1-20.

Prather, R. W., & Alibali, M. W. (2008). Understanding and using principles of arithmetic: Operations involving negative numbers. *Cognitive Science, 32*(2), 445-457.

Pratt, D. D. (1998). *Five perspectives on teaching in adult and higher education*: Malabar, Fla. : Krieger Pub. Co. Original ed.

Prestopnik, N., & Crowston, K. (2012). *Purposeful gaming & socio-computational systems: a citizen science design case.* Paper presented at the Proceedings of the 17th ACM international conference on Supporting group work.

Price, M., Carroll, J., O'Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: a critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education, 36*(4), 479-492.

Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *MIS quarterly*, 731-755.

Privitera, G. J. (2011). *Statistics for the behavioral sciences*: Sage.

Randolph, J. J. (2009). A guide to writing the dissertation literature review. *Practical Assessment, Research & Evaluation, 14*(13), 1-13.

Redmond, P. (2014). Reflection as an indicator of cognitive presence. *E-Learning and Digital Media, 11*(1), 46-58.

Riff, D., Lacy, S., & Fico, F. (2014). *Analyzing Media Messages. Using Quantitative Content Analysis in Research*: Hoboken : Taylor and Francis 3rd ed.

Ritchie, J., Lewis, J., McNaughton Nicholls, C., & Ormston, R. (2014). *Qualitative research practice : a guide for social science students and researchers*: London : SAGE Publications Ltd, 2014.

Second edition.

Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (2013). *Qualitative research practice: A guide for social science students and researchers*: sage.

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences, 4*(2), 155-169.

Roberts, J. M. (2014). Critical realism, dialectics, and qualitative research methods. *Journal for the Theory of Social Behaviour, 44*(1), 1-23.

Robinson, S. (2015). *A tutorial on conceptual modeling for simulation.* Paper presented at the Proceedings of the 2015 Winter Simulation Conference.

Rønning, F. (2017). Influence of computer-aided assessment on ways of working with mathematics.

Teaching Mathematics and Its Applications: International Journal of the IMA, 36(2), 94-107.

Rose, D. H., Robinson, K. H., Hall, T. E., Coyne, P., Jackson, R. M., Stahl, W. M., & Wilcauskas, S. L. (2018). Accurate and Informative for All: Universal Design for Learning (UDL) and the Future of Assessment *Handbook of Accessible Instruction and Testing Practices* (pp. 167-180): Springer.

Rose, S. (2015). *Management research. applying the principles*: New York : Routledge
1st ed.

Rowe, F. (2014). What literature review is not: diversity, boundaries and recommendations: Taylor & Francis.

Sagitova, R. (2016). *Climate Change and GHG Emissions Reporting in Russia: practice and perceptions.* Heriot-Watt University.

Salmona, M., & Smith, T. (2014). Qualitative Research in Finance. *Australian Journal of Management, 39*(1).

Sangwin, C. (2012). *Computer aided assessment of mathematics using STACK*. Sangwin, C. (2013). *Computer aided assessment of mathematics*: OUP Oxford.

Sangwin, C. (2015a). *Computer aided assessment of mathematics using STACK.* Paper presented at the Selected Regular Lectures from the 12th International Congress on Mathematical Education.

Sangwin, C. (2015b). Computer Aided Assessment of Mathematics Using STACK. In J. S. Cho (Ed.), *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 695-713). Cham: Springer International Publishing.

Sangwin, C., Cazes, C., Lee, A., & Wong, K. L. (2009). Micro-Level Automatic Assessment Supported by Digital Technologies. *13*, 227-250. doi:10.1007/978-1-4419-0146-0_10

Sangwin, C. J. (2006). Mathematical question spaces.

Sangwin, C. J., & Köcher, N. (2016). Automation of mathematics examinations. *Computers & Education, 94*, 215-227.

Saunders, M., Lewis, P., & Thornhill, A. (2009). Understanding research philosophies and approaches. *Research methods for business students, 4*, 106-135.

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (7th ed.). Harlow, UK: Pearson.

Sayer, A. (2000). *Realism and social science*: Sage.

Schermann, M., Böhmann, T., & Krcmar, H. (2009). Explicating design theories with conceptual models: Towards a theoretical role of reference models *Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik* (pp. 175-194): Springer.

Schön, D. A. (1983). The reflective practitionerhow professionals think in action.

Scott, D. (2007). Resolving the quantitative–qualitative dilemma: a critical realist approach. *International Journal of Research & Method in Education, 30*(1), 3-17.

Seale, C., & Silverman, D. (1997). Ensuring rigour in qualitative research. *The European Journal of Public Health, 7*(4), 379-384.

Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research.

Management Information Systems Quarterly, 35(1), 37-56.

Shanks, G., Tansley, E., & Weber, R. (2003). Using ontology to validate conceptual models. *Communications of the ACM, 46*(10), 85-89.

Shedlock, K., Vos, M., & Chard, S. (2016). Design Science Research: Building evaluation into the construction of indigenous cultural artefacts in New Zealand.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational leadership, 63*(3), 66- 70.

Siddaway, A. (2014). What is a systematic literature review and how do I do one. *University of Stirling*(I), 1.

Silverman, D., & Marvasti, A. B. (2008). *Doing qualitative research : a comprehensive guide*: Los Angeles : SAGE Publications

2nd ed.

Sim, G., Holifield, P., & Brown, M. (2004). Implementation of computer assisted assessment: lessons from the literature. *ALT-J, 12*(3), 215-229.

Simkin, M. G., & Kuechler, W. L. (2005). Multiple-Choice Tests and Student Understanding: What Is the Connection? *Decision Sciences Journal of Innovative Education, 3*(1), 73-98.

Singh, U. G., & de Villiers, M. R. (2017). An Evaluation Framework and Instrument for Evaluating e- Assessment Tools. *International Review of Research in Open & Distance Learning, 18*(6), 164- 185.

Sobh, R., & Perry, C. (2006). Research design and data analysis in realism research. *European Journal of marketing, 40*(11/12), 1194-1209.

Soiferman, L. K. (2010). Compare and Contrast Inductive and Deductive Research Approaches. *Online Submission*.

Solving Systems of Linear Equations; Row Reduction - HMC Calculus. Solving Systems of Linear Equations; Row Reduction. Retrieved from https://www.math.hmc.edu/calculus/tutorials/linearsystems/

Sonnleitner, P., Keller, U., Martin, R., & Brunner, M. (2013). Students' complex problem-solving abilities: Their structure and relations to reasoning ability and educational success.

Intelligence, 41(5), 289-305. doi:10.1016/j.intell.2013.05.002

Stickley, T. (2006). Should service user involvement be consigned to history? A critical realist perspective. *Journal of Psychiatric and Mental Health Nursing, 13*(5), 570-577.

Suskie, L. (2010). Assessing Student Learning: A Common Sense Guide ,(2009): San Francisco: Jossey- Bass.

Szopinski, D., Schoormann, T., & Kundisch, D. (2019). BECAUSE YOUR TAXONOMY IS WORTH IT: TOWARDS A FRAMEWORK FOR TAXONOMY EVALUATION.

Tashakkori, A., & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research*: Sage.

Taylor, R. R. (2017). *Kielhofner's Research in Occupational Therapy: Methods of Inquiry for Enhancing Practice*: FA Davis.

Thuan, N. H., Drechsler, A., & Antunes, P. (2019). Construction of design science research questions. *Commun. Assoc. Inf. Syst, 44*(1), 20.

Tilly, C. (2008). *Explaining social processes*: Boulder : Paradigm Publishers.

Vaishnavi, V., Kuechler, W., & Petter, S. (2017). Design Science Research in Information Systems.

Vaishnavi , V., Kuechler, W., & Petter, S. (2017). Design science research in information systems.

Retrieved from http://www.desrist.org/design-research-ininformation-systems/

Vaishnavi, V. K., & Kuechler, W. (2015). *Design science research methods and patterns: innovating information and communication technology*: CRC Press.

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice, 6*(5), 100.

Veloski, J. J., Rabinowitz, H. K., Robeson, M. R., & Young, P. R. (1999). Patients don't present with five choices: an alternative to multiple-choice tests in assessing physicians' competence. *Academic medicine, 74*(5), 539-546.

Venable, J. (2006). *A framework for design science research activities.* Paper presented at the Emerging Trends and Challenges in Information Technology Management:

Proceedings of the 2006 Information Resource Management Association Conference.

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European Journal of Information Systems, 25*(1), 77-89.

Venable, J. R. (2010). *Design science research post hevner et al.: criteria, standards, guidelines, and expectations.* Paper presented at the International Conference on Design Science Research in Information Systems.

Venkatesan, V. P., & Kuppuswami, S. (2008). ARMMS-Architecture Reference Model for Multilingual Software. *Journal of Convergence Information Technology, 3*(2), 74-81.

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS quarterly, 37*(1).

Vlassis, J. (2004). Making sense of the minus sign or becoming flexible in 'negativity'. *Learning and instruction, 14*(5), 469-484.

Vogt. (2005). *Dictionary of Statistics & Methodology. [electronic resource] : A Nontechnical Guide for the Social Sciences (Third Edition)*: [S.l.] : Sage Publications, Inc, 2005.

Vom Brocke, J., & Buddendick, C. (2006). *Reusable conceptual models–requirements based on the design science research paradigm.* Paper presented at the Proceedings of the First International Conference on Design Science Research in Information Systems and Technology (DESRIST).

Wacker, J. G. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of operations management, 16*(4), 361-385.

Walker, P., Gwynllyw, D. R., & Henderson, K. L. (2015). Diagnosing student errors in e-assessment questions. *Teaching Mathematics and its Applications: An International Journal of the IMA, 34*(3), 160-170.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an information system design theory for vigilant EIS. *Information systems research, 3*(1), 36-59.

Wand, Y., & Weber, R. (2002). Research commentary: information systems and conceptual modeling- a research Agenda. *Information systems research, 13*, 363-376.

Weber, S. (2010). *Design Science Research: Paradigm or Approach?* Paper presented at the AMCIS. White, A. (2019). A Discussion and Students' Thoughts on the Assessment Methods used at the Tertiary Level.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3-14. doi:10.1016/j.stueduc.2011.03.001

Wiliam, D. (2013). Assessment: The bridge between teaching and learning. *Voices from the Middle, 21*(2), 15.

Wong, G., Greenhalgh, T., Westhorp, G., & Pawson, R. (2012). Realist methods in medical education research: what are they and what can they contribute? *Medical Education, 46*(1), 89-96. doi:10.1111/j.1365-2923.2011.04045.x

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence, 40*(1), 1-14. doi:10.1016/j.intell.2011.11.003

Yeo, R. K., & Marquardt, M. J. (2012). Complex problem solving through action learning: implications for human resource development. *International Journal of Human Resources Development and Management, 12*(4), 258-273.

Zachariadis, M., Scott, S. V., & Barrett, M. I. (2010). Exploring critical realism as the theoretical foundation of mixed-method research: evidence from the economies for IS innovations.

Appendix A

Questionnaire

The questionnaire structure reflects the steps that occur in sequence to analyse the given examinations' questions in the student scripts Table A.1 shows 32 questions in this questionnaire. Questions one and two are the case identification information. Since each student script is kept in a separate record within a group, a numerical value is given to the script for identification purposes. This information is covered by the first question. The second question gathers the data that identifies the group that each student script belongs to.

Question three collects the data that presents whether the student has miscopied the original question or not. The fourth question records whether equations have been swapped in the student workings. Questions five, six and seven gather the data that shows the order in which elements in the matrix are eliminated from the matrix, for example which element has been eliminated in the Gaussian elimination process.

In the student workings, any of the elements may be selected for elimination from the matrix. Therefore, in the design of scripts' analysis, the possibility of any of the nine elements of matrix- in 4x4 matrices- selection for elimination had been taken into consideration. To accommodate for this scenario, the questionnaire included questions for all terms in the matrix. Questions 8 to 18 cover the data (solutions strategies to eliminate elements and the occurrence of associated errors) to be collected for a11 element from the list of possible elements that might be eliminated from the matrix. The data to be analysed has been designed to accommodate for this scenario; the tasks to perform actions that are requested by questions 8 to 18 has been extended to cover other elements, which are a12, a13, a21, a22, a23, a31, a32 and a33. However, from the data that has been produced from the qualitative phase findings, only three elements, which are a21, a31 and a32 are involved in elements elimination. This means the data that belongs to the elimination actions of a11, a12, a13, a22, a23 and a33 will be blank.

Question 19 is used to record the equation that is left with one unknown and its value, after the completion of the element's elimination process. To collect the data about the first unknown that its value is computed for is displayed in question 20. Question 21 collects the type of solution strategy that student has used to compute for the value of the first unknown that is identified in question 20.

Questions 22 to 29 involve collecting data (unknown values computation solution strategies and the associated errors) regarding z variable. The same information that is gathered by questions 22 to 29 are used to compute for x and y variables; but separate variables are used to gather data for x and y. Questions 30 and 31 are used to gather data about the occurrences of combining steps when performing elements' elimination and unknown value computation actions. Finally, the task of question 32 is to collect the actual mark that has been assigned to the student workings that are shown in the script.

Table A.1 Questionnaire

Questions	Questions Wordings
1	What is the scripts number (Script-ID) within the group?
2	Which group (script-Group) does the script belong to Group 1 = 20111 Group 2 = 20112 Group 3 = 2012 Group 4 = 2013
3	Has the original question miscopying (q-Cpy-Err) occurred? Yes = 1 if no miscopying, just leave the variable's value blank
4	Which equations (swp-Eqn) have been swapped? Eq1-with-Eq2 = 1 Eq1-with-Eq3 = 2

	Eq2-with-Eq3 = 3 If no swapping action has not happened, leave the variable value blank
5	Which element (frt-Elm-V) has been eliminated first from a general 3x3 matrix? a11 =1 a12= 2 a13 = 3 a21= 4 a22=5 a23=6 a31=7 a32=8 a33=9 If a variable has not been eliminated, leave the variable value blank
6	Which element (snd-Elm-V) has been eliminated second from a general 3x3 matrix? a11 =1 a12= 2 a13 = 3 a21= 4 a22=5 a23=6 a31=7 a32=8 a33=9 If a variable has not been eliminated, leave the variable value blank
7	Which element (trd-Elm-V) has been eliminated third from a general 3x3 matrix? a11 =1 a12= 2 a13 = 3 a21= 4 a22=5 a23=6 a31=7 a32=8 a33=9
8	Which elimination method has been used to eliminate an (v1-Elm-Mth) element to form an upper triangular matrix in a 3x3 matrix? Correct expected =1 Correct acceptable =2 Correct expected no desired output = 3 Correct acceptable no desired output =4 Correct expected with info = 5 Correct acceptable with info = 6 Incorrect unacceptable no info = 7 Incorrect unacceptable no info = 8 Incorrect unacceptable with info accepting solution = 9 Correct reintroduces solved problems = 10 Unnecessary Work with error and-or info = 11 Unnecessary Work no error or info = 12 If the element elimination method has not been provided, leave the variable value blank
9	Has the student provided correct or acceptable method (v1-mth-Not-Ap) to eliminate a11 element without applying the method? Yes = 1 Provided incorrect method = 2 If a variable elimination method has not been provided to eliminate the a11 element, leave the variable value blank

10	Has an arithmetic error (v1-Arh-Mth-Apl-Err) occurred during a11 element elimination application? Yes once = 1 Yes multiple = 2 If an arithmetic error has not been detected, leave the variable value blank
11	Has a sign error (V1-Sgn-Mth-Apl-Err) occurred during a11 element elimination method application? (variable name v1 refers to a11 element in equation 1) Yes once = 1 Yes multiple = 2 If a sign error has not been detected, leave the variable value blank
12	Has a double negative error (V1-Dng-Mth-Apl-Err) occurred during a11 element elimination method application? (variable name v1 refers to a11 element in equation 1) Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
13	Has an Eliminated variable only operation error (v1-EVOO-Math-Apl-Err) occurred during a11 element elimination method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
14	Has a purposeful variable elimination error (v1-PVEE-Mth-Apl-Err) occurred during a11 element elimination method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
15	Has a miscopying error (v1-msp-Mth-Apl-Err) occurred during a11 element elimination method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
16	Has a follow through error (v1folo-Through-Err) occurred, during a11 element elimination method application, present? (To record the presence of follow through error in the current step's workings.) Yes once = 1

	Yes multiple = 2 (during the a11 element elimination process more than one step could be used to eliminate a11; if more than one step produces incorrect answer, more than one follow through error might happen) If this type of error has not been detected, leave the variable value blank
17	Has the ignored or skipped part of operation (Ignrd-Mth-Ap-Err) error occurred during a11 element elimination method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
18	Has an others error type (v1-Otr-Mth-Apl-Err) occurred during a11 element elimination method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
19	Which equation (eqn-With-1-mVar) is left with one unknown after the completion of elements elimination actions? Equation 1 =1 Equation 2 =2 Equation 3 =3 More than one equation = 4 If there is no equation with one variable only, leave the variable value blank
20	Which unknown's value has been computed for first? X=1 Y=2 Z=3 If there no answer for this question, leave the variable value blank
21	Which computation method (frt-Var-Val-Cmtd-Mth) has been used in finding the value of the first unknown that its value is computed for? Correct expected = 1 Correct acceptable = 2 Incorrect = 3 If the method has not been given, leave the variable value blank.

22	Which computation method (z-Val-Cmtd-Mth) has been used in finding the Z value? Correct expected = 1 Correct acceptable = 2 Incorrect = 3 If the method has not been given, leave the variable value blank.

23	Has an arithmetic error (z-Arh-Val-Cmtd-Mth-Apl-Err) occurred during the z value computation method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
24	Has a sign error (z-Sgn-Val-Cmtd-Mth-Apl-Err) occurred during the z value computation method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
25	Has a double negative error (z-Dng-Val-Cmtd-Mth-Apl-Err) occurred during the z value computation method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
26	Has a miscopying error (z-msp-Val-Cmtd-Mth-Apl-Err) occurred during the z value computation method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank
27	Has an equation rearrangement error (z-ER-Val-Cmtd-Mth-Apl-Err) occurred during the z value computation method application?
28	Has a follow through (z-folo-Through-Err) error, during the z value computation method application, present? Yes once = 1 Yes multiple = 2 (During the z value computation, there could be more than one steps; if more than one step involves in producing incorrect answer, more than one follow through error might happen) If this type of error has not been detected, leave the variable value blank
29	Has an others error type (v1-Otr-Mth-Apl-Err) occurred during the z value computation method application? Yes once = 1 Yes multiple = 2 If this type of error has not been detected, leave the variable value blank

30	Has steps combination (mVar-Elm-Stp-Cmb), during element elimination process, action occurred? Yes once = 1 Yes multiple = 2 If this type of action has not been detected, leave the variable value blank
31	Has steps combination (mVar-Val-Stp-Cmb), during unknowns' values computation process, action occurred? Yes once = 1 Yes multiple = 2 If this type of action has not been detected, leave the variable value blank
32	What mark has been awarded to this working? Mark1 = 0.17 Mark2 = 0.2 Mark3 =0.25 Mark4 =0.3 Mark5 =0.33 Mark6 =0.4 Mark7 =0.5 Mark8 =0.6 Mark9 =0.7 Mark10 =0.75 Mark11 =0.8 Mark12 =0.9 Mark13 = 1.0

Appendix B

Categories and subcategories of solution strategies and errors

This appendix contains further details concerning the categories and subcategories that have been formed and reported in section 5.4. In this appendix, there are two types of information that will be displayed in tables. This two information was created from the process of categorising the result of the qualitative content analysis of student workings. Table B.1 to Table B.8 display information about subcategories of solution strategy, while, Table B.9 toTable B.19 contain information about student errors. The meanings of the columns in Table B.1 to Table B.8 have been given in Table 5.1. The 'Error' column in these tables shows the relationships between solution strategy subcategories and the error subcategories.

Table B.1 Correct expected and correct acceptable solution strategies subcategories

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
1	Correct expected	✓	×	✓	✓
2	Correct acceptable	✓	×	✓	✓
Explanations The correct application of these subcategories results in producing the desired outputs that do not contain any error. The outputs from the two subcategories are accepted as the valid inputs in the following steps and these outputs are awarded full marks.					

Table B.2 Correct expected without desired output and correct acceptable without desired output

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
3	Correct expected without desired output	N/A	N/A	N/A	✓X
4	Correct acceptable without desired output	N/A	N/A	N/A	✓X

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
Explanations These subcategories refer to the expected or acceptable type solution strategies that students display without applying the strategies to produce the desired output. Therefore, the desired output, error and valid input columns are filled with the N/A value to express that the displayed solutions strategies have no effect on these columns' values. However, the awarded column value is both tick and X to show partial mark is awarded; this is to indicate the knowledge of only correct or acceptable solution strategy, to perform a specific task, warrants a partial mark.					

Table B.3 Correct expected with extractable info and acceptable with extractable info

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
5	Correct expected with extractable info	✕	✓	✓✕	✓✕
6	Correct acceptable with extractable info	✕	✓	✓✕	✓✕
Explanations These refer to solution strategies which are applied to perform the task; however, as the columns display, the desired output is incorrect indicating there is/are 'Errors', as shown in the error column. In the 'valid input' column, tick and x, are both shown to indicate that the output of the process might be used as valid input in performing the next step of the task. The awarded column, similarly, shows tick and x to indicate that the student working is awarded a partial mark for providing a correct solution strategy and in some cases extra marks for the workings showing some of the knowledge during the application of the solutions strategy.					

Table B.4 Incorrect unacceptable/unrelated with no info

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
7	Incorrect unacceptable/unrelated with no info	N/A	N/A	N/A	N/A
Explanations This subcategory contains the incorrect solution strategy that does not contain any assessable workings. Such work may provide insight into student understanding even if formally is irrelevant to the working or solution strategy. All the columns values are					

filled with N/A. However, regarding the mark awarded to the workings, this should be 0 for not demonstrating any knowledge.

Table B.5 Incorrect unacceptable/unrelated with info

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
8	Incorrect unacceptable/unrelated with info	✕	✓	✕	✕

Explanations

This is similar to the subcategory 7. The examining of the columns' values, however, shows the difference between the two subcategories. In subcategory 8, the columns are filled with Xs except the error column's value which is now ticked to indicate there is/are errors in the workings. Thus, the purpose of gathering the information that occurs due to the application of the subcategory 8, 'incorrect solution strategy', is to use the collected error in the analysis of the student workings.

Table B.6 Incorrect unacceptable/unrelated with info

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
9	Incorrect unacceptable/unrelated with info and accepting the solution produced	✕	✓	✓	✕

Explanations

This is similar to the subcategory 8; the only difference is that the valid input column has a tick indicating that the desired output value is accepted as the valid input in the following step's task.

Table B.7 Correct solution strategies that reintroduces previously solved problems

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
10	Correct solution strategies that reintroduces previously solved problems	✓	✓	✓	✓X

Explanations

In this subcategory, the working shows that the solution strategy that is used is correct to perform the current step task producing the 'desired output' for this particular step; however, this type of solution strategy reintroduces a previously solved part of the overall question. Therefore, the 'purposeful type of error' is recorded against the application of this solution strategy type. The valid input column contains a tick to show that this solution strategy type allows the students to use the 'desired output' as the valid input in the next step; the value of the awarded column, similarly, shows that the student workings using this solution strategy justifies awarding them partial mark.

Table B.8 Unnecessary work with error and/or info and unnecessary work with no error or info

Entities' numbers	**Solutions strategies Categories and subcategories names**	**Desired Outputs**	**Errors**	**Valid inputs**	**Awarded**
11	Excessive/Unnecessary Work with error or/and info	×	✓	×	×
12	Excessive/Unnecessary Work with no error or info	×	×	×	×

Explanations

Subcategory 11 could have the features of all subcategories 1 to 10, excluding subcategory 7, at different times. This means that the solution strategy of the subcategory 11 could be the type of the strategy that might be correct expected, correct acceptable or incorrect type that includes error or errors. Subcategories 11 and 12 represent excessive works that are not related to solving of the questions. The analysis of the subcategory 11 shows that the columns desired output, valid input and awarded are filled with the value of x. The value x shows that the desired output is not accepted as the valid current step output. The valid input value x also indicates that the desired output is not accepted as the valid input in the next step. The awarded value x shows that this working is not awarded any mark. The types of errors in the solution strategy that is represented by subcategory 11 could be any type of error that is uncovered in subcategories 3-10 solution strategies, except the subcategory 7. The solution strategies represented by subcategory 1 and 2 do not contain any error and the solution strategy that is represented by subcategory 11 must contain any types of error that might be shown in the other solution strategies represented by subcategory 3 – 10. Subcategory 12 only differ from subcategory 11 in that it does not introduce any error; and consequently, it does not affect the overall solution of a question.

Table B.9 Arithmetic error subcategories entity, meaning and effects

1- Arithmetic error	
Definition:	The arithmetic error is the error which is formed because of the incorrect mathematical computation excluding negative sign, double negative sign and purposeful errors.
Source:	human (incorrect application of solution strategy or introducing other errors)
Cause of Error:	When a term elimination and/or values computation tasks are performed.
Consequence of Error on Other Tasks Performing:	The output of the current step, which has produced the arithmetic error, is accepted as the valid input in the subsequent operations with the acknowledgement of the follow through error
Consequence of error on awarding mark:	Awarded partial mark
Comment:	It is necessary to search and identify arithmetic errors to reveal their effect on the steps results.

Table B.10 Follow through error subcategories entity, meaning and effects

2- Follow through error	
Definition:	The “follow through error” is an error that might be single or multiple and it occurs in previous steps; and transmits to the following steps. The presence of a follow through error in those subsequent steps changes results of those steps but should have no effect in awarding marks to these steps’ workings.
Source:	None (the error has been identified in previous step(s))
Cause of Error:	None, it is the error that has been transmitted from the output of the previous step’s incorrect output and it is used as the input in the current step.
Consequence of error on other tasks performing:	The output of the current step involves the usage of the previous step or steps output(s). Therefore, the usage of the incorrect answer in the current step’s tasks produces the incorrect answer even if the current step’s operation is correctly performed. Thus, the subsequent steps operations answers are incorrect too; however, these incorrect answers might receive full marks, provided their workings are correct.

Consequence of error on awarding mark:	The mark awarding is not dependent on the follow through error, but on the workings that has been shown in the current step.
Comment:	It is necessary to acknowledge the presence of the follow through error in the following steps; because its presence in subsequent steps produces different answers than the original expected steps' results.

Table B.11 Equation Rearrangement error subcategories entity, meaning and effects

3- Solution Strategies Modification	
Definition:	The 'solution strategies modification error' arises from luck of knowledge in changing the sing of a subject or signs of subjects. In this case the error occurs due to the equation rearrangement, although the equation rearrangement by itself has no effect mathematically. This error can be considered as an instance of the 'in-working-miscopying error' type. However, due to the reason causing this error it has been categorised separately.
Source:	Solution Strategy
Cause of Error:	Lack of knowledge (when rearranging equations)
Consequence of error on other tasks performing:	The execution of the modified solution strategy, due to equation rearrangement, produces a different answer than the answer that is produced using the correct equation rearrangement method. Therefore, the output of this current answer, if used, changes the output of the subsequent step(s)' answers.
Consequence of error on awarding mark:	The error that is made when rearranging equations is similar to the act of changing solutions strategies; however, since the initial equation to perform the task(s) is correct, partial mark is applicable to the workings.
Comment:	It is required to search for the "incorrect equations rearrangement error' that causes the modification of the expected step result.

Table B.12 Purposeful error subcategory entity, meaning and effects

4- Purposeful error	
Definition:	The ‘purposeful error’ is the error that students intentionally make to produce the intended correct result. This occurs when using the incorrect solution strategy that produces a partially correct answer, but also introduces a partly incorrect solution. However, students ignore the incorrect part of the answer and use the correct part as the correct solution.
Source:	Solution strategy and human (incorrect application of solution strategy or introducing other errors)
Cause of Error:	Lack of selecting correct solution strategy (especially when eliminating terms from equations)
Consequence of error on other tasks performing:	The step that yields the purposeful error produces an incorrect answer; this answer can be used as a valid input in the subsequent steps, however the follow through errors must be acknowledged.
Consequence of error on awarding mark:	Since the purposeful error is purposefully made to display the correct knowledge, a partial mark is awarded to the workings for producing the required result with the current working containing incorrect solution strategy.
Comment:	It is required to check only changes of the solution strategy and/or changing the results of the strategy to produce correct current answer using the incorrect solution strategies.

Table B.13 Others error subcategory entity, meaning and effects

5- Others error	
Definition:	This is an unexplainable type of error within a step affecting the operation in the step.
Source:	Solution strategy and human (incorrect application of solution strategy or introducing other errors)
Cause of error:	The unexplainable type of error can occur at any stage in solving questions
Consequence of error on other tasks performing:	If an output of any operation that contains the unexplainable error is used in the subsequent steps, it can cause any types of error. The output of this current solution strategy can be used in the subsequent operations; however, the acknowledgement of the presence of the follow through error is required.
Consequence of error on	No mark is awarded to the result that contains these errors.

awarding mark:	
Comment:	It is required to identify and report the presence of this type of error.

Table B.14 Negative error subcategory entity, meaning and effects

6- Negative Sign Error	
Definition:	This is the error that occurs when students ignore or forget the negativity of numbers when applying the correct solution strategy to perform tasks and produce incorrect answer.
Source:	Solution strategy and human (incorrect application of solution strategy or introducing other errors)
Cause of Error:	This error is caused either by a human error (slip) which is not related to the knowledge of the problem solving or by incorrect usage of minus signs.
of error on other tasks performing :	The output of the current step is accepted as the valid input in the subsequent steps provided no other errors that prevent the output to be accepted as the valid input.
Consequence of error on awarding mark:	Partial mark awarding is possible.
Comment:	It is required to check and identify the presence of the negative sign error, because it causes the production of incorrect steps answers introducing the follow through error in the following steps workings and it also might cause the creation of the purposeful error, or an arithmetic error, or the creation of a double minus error.

Table B.15 Double Negative Sign Error subcategory entity, meaning and effects

7- Double Negative Sign Error	
Definition:	The double negative sign error is the error that is not a negative sign error but is caused by the presence of the subtraction operation using negative numbers in which the problem solvers fail to convert the double negative sign to the positive sign.
Source:	Solution Strategy and human (incorrect application of solution strategy or introducing other errors)

Cause of Error:	Lack of Knowledge or human error
Consequence of error on other tasks performing:	The output of the current step is accepted as the valid input in the next steps provided that no other errors that prevent the current output to be accepted as the valid input
Consequence of error on awarding mark:	Partial mark awarding is possible.
Comment:	It is required to check and identify the presence of the double minus sign error, because it causes the production of incorrect steps answers and it also causes the introduction of the follow through error or the creation of purposeful error or arithmetic error.

Table B.16 Ignored operation error subcategory entity, meaning and effects

8- Ignored/Skipped Operation error	
Definition:	The 'ignored operation error' is the error that occurs when students ignore a mathematical operation in a part of a task or process that is expected to achieve an aim, for example term's elimination from an equation, producing an incorrect answer. It resembles the purposeful error, but it is different.
Source:	Solution strategy and human (incorrect application of solution strategy or introducing other errors)
Cause of Error:	The usage of the incorrect solution strategy and/or the improper application of the correct solution strategy.
Consequence of error on other tasks performing:	The output of current incorrect answer is used in the following step as the valid input; however, the introduction of follow through error in the subsequent steps must be acknowledged
Consequence of error on awarding mark:	Partial mark awarding is possible.
Comment:	It is necessary to check and identify the occurrence of the ignored operation error.

Table B.17 Eliminated variable only operation error subcategory entity, meaning and effects

9- Eliminated Variable Only Operation	
Definition:	This error occurs when the correct solution strategy is provided to eliminate a particular term from an equation; however, the column in the equation that contains the term to be eliminated is the only part in the equation that is affected by the solution strategy application. The

	other columns, in the equation, remain unchanged. The solution strategy is selectively applied to yield only the desired result.
Source:	Solution strategy and human (incorrect application of solution strategy or introducing other errors)
Cause of Error:	Human purposeful error
Consequence of error on other tasks performing:	The output of the current step is accepted as the valid input in the subsequent steps if there are no other errors. If errors are present then, the output cannot be accepted as the valid input for the subsequent steps. Since the desired output is partially incorrect, the follow through error is acknowledged in the following workings.
Consequence of error on awarding mark:	The partial mark is awarded for using the correct solution strategy, although the operation is incorrectly applied. This is similar to the arithmetic error, however, the reason that causes the error is different.
Comment:	It is required to check, identify and report the occurrence of this error, because this action changes the expected result and has consequences on the following steps tasks.

Table B.18 Original question miscopying error subcategory entity, meaning and effects

10- Original Question miscopying	
Definition:	It is the error in which the original examination question is altered. It may cause the usage of the different solution strategies that yields different answers; consequently, different types of errors might result from the miscopied question.
Source:	Human
Cause of Error:	Human error
Consequence of error on other tasks performing:	Not applicable to this type of error
Consequence of error on awarding mark:	None; awarding mark is dependent on any error that occurs in the workings not related to the miscopying of the original question.
Comment:	It is required to check and identify the occurrence of the ‘original question miscopying’ error, because this condition changes the status of the original question solution strategies and associated errors that might happen.

Table B.19 In-working miscopying subcategory entity, meaning and effects

11- In-working miscopying	
Definition:	It is the error that occurs due to miscopying any section of the working: numbers, symbols, coefficients, and mathematical signs of these entities during the solution. This error has similar effects as the original question miscopying error, although the change in the question's solution is only applies to part of the solution.
Source:	Human (This type of error is not related to selecting or using any types of solution strategies)
Cause of Error:	Human error
Consequence of error on other tasks performing:	Accept the output of the current step as the valid input in the subsequent steps with the acknowledgement of the follow through error; however, the acceptance of the current step answer as the valid input in the subsequent steps is dependent on the absence of other types of errors that cause the current step output as an invalid input.
Consequence of error on awarding mark:	The presence of the miscopying error is ignored, and awarding is dependent on the presence of other types of errors
Comment:	It is necessary to check and identify this error to take a proper response.