Effects of a mathematics teaching strategy based on distributed interleaved practice on procedural and conceptual knowledge.

Cindy Carelse

3446346

A thesis submitted in fulfilment of the requirements for the degree of master's in education in the School of Science and Mathematics Education (SSME), University of the Western Cape.

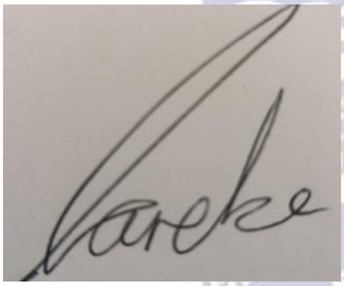
10 October 2022

Supervisor: Dr. B May

Co-supervisor: Prof C. Julie

Declaration

I, Cindy Carelse, declare that the contents of this thesis represent my unaided work and that the thesis has not previously been submitted for examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of participants.



Signed: ________________ Date: _27 October2022__

Acknowledgments:

I would like to thank the following people who have helped me undertake this research:

My supervisor Dr B May., for his enthusiasm for the project, for his support, encouragement and patience.

The principal Mr C Arries of the primary school who was so generous in allowing me to complete my research.

My partner Dr L Twala – I simply couldn't have done this without you, special thanks.

And to my parents Clive and Rachel Carelse, who set me off on the road to this M.Ed. a long time ago, I owe you everything.

Contents

Abstract

This study investigated the effects a teaching strategy based on interleaved distributed practice had on learners in three grade 6 classes' procedural and conceptual knowledge within a, previously disadvantaged, Cape Town Primary School. An inference one can draw from the pace setters of the South African Curriculum (CAPS) is that mass practice is the suggested practice type. Mass practice is a teaching strategy where learners would practice problems requiring one or more specific skill(s) immediately after the presented lesson on that skill(s). The aforementioned practice allowed learners to predict the type of problems they would encounter in these activities. The interleaved distributed practice presents an alternative teaching strategy to massed practice. In this new practice, activities were organized in a manner that includes more than one skill and is presented in an interwoven manner with other related skills. This strategy allows learners to see each problem within activities as an individual instead of one continuous skill being practiced. Therefore, they would subsequently learn how to solve these problems on individual bases using prior knowledge and skills gained, as opposed to skills learned only in the preceding lesson. The constant revising process within the interleaved distributed practice would aid learners' procedural and conceptual knowledge which in turn could aid learners' procedural flexibility as learners would then be able to build the skills in order to adapt and apply their skills to a variety of problems. This study made use of a quasi-experimental research approach. The teaching intervention (interleaved distributed practice) were used with two classes in an alternating manner while a third class (used as a control group) were not exposed to the intervention. Class groups A and B were used as experimental and control groups in an alternating manner. A revised taxonomy table was utilized to determine how the different knowledge components (procedural and conceptual) were affected. Our findings show that over time interleaved distributed practice increases memory retention as well as procedural and conceptual understanding. Whereas when a massed practice strategy is utilized knowledge retention, in terms of procedural and conceptual understanding decreases over time.

KEYWORDS:

Mass practice; knowledge retention; interleaved and distributed practice; procedural and conceptual knowledge

Chapter 1: Introduction and background

1. Introduction and Background

The South African education system is viewed by many education practitioners as a structure where the following statement applies: "*South Africa is significantly underperforming in education in general, particularly mathematics teaching and learning*" (Motshekga, 2016, p1). Using the Trends in Mathematics and Science Studies (TIMMS) results from 2002 to 2011 to substantiate, Feza (2014) agrees with the then South African minister of basic education. Feza (2014) noted that although there was an improvement in the results of South African learners they were still regarded as underachieving in comparison to most other countries. TIMMS is an international data study that compares the achievement of learners in Mathematics and Science. South Africa was a participant in the TIMMS study during the years 1995,1999,2003,2011 and 2015 (Reddy, Visser, Winnaar, Arends, Juan, Prinsloo & Isdale, 2016). The TIMMS collects data from learners in grades 4, 8, and 12. A TIMMS study is completed every four years (https://nces.ed.gov/timss/).

Many similar schools, in the area where I teach, echo the findings of TIMMS. This can be seen in the results of the Systemic tests. Since 2014 the Western Cape Education Department (WCED) implemented systemic testing. Systemic tests are written by all Western Cape learners at the end of three phases of schooling. For example, in grade 3 (as this is the end of the foundation phase), in grade 6 (as this is the end of the intermediate phase), and in grade 9 (as this is the end of the senior phase). According to the WCED (2016), these test results are used by the WCED in order identify where improvement can be made.

My school (a quintile 3 school) located on the Cape Flats has been taking these tests and based on the grade 6 results (the end of the intermediate phase) the school has been labelled as an underperforming school. According to the WCED (2014), the quintile label is one way the WCED categorizes schools for purposes such as resources and financial allocation. Quintile 1 would be seen as the poorest and quintile 5 the least poor.

The 2018 Systemic test results show that most learners at my school are underperforming. Only 33.5% of learners managed to achieve a pass percentage in mathematics (A Hendricks, personal communication, March 18, 2019). Ms. A Hendricks is the school secretary responsible for distributing the summaries and conclusion of the systemic results to the school staff. It is concerning that only 33.5% of learners can achieve 50% and above in this mathematics test.

The minimum pass requirement is 50% for mathematics in the intermediate phase (grades 4-6), in South Africa. These results are concerning since the mathematics content covered in the intermediate phase lays the foundation for all mathematics concepts in future years, such as algebra, geometry, and trigonometry.

My motivation for this study was to contribute to learner proficiency and understanding of mathematics by considering different interventions. It is common knowledge that proficiency in mathematics can be enhanced through practice. I intended to implement an intervention based on a different way of practice with the objective to improve conceptual and procedural knowledge in mathematics.

Improvement of conceptual and procedural knowledge of learners in mathematics is important as the mathematical knowledge of the intermediate phase (grades 4-6) lays the foundation for mathematics learning in the senior phase. If learners move into the senior phase with a good understanding, they will more likely gain a positive attitude toward mathematics and have better attainment. Whereas, if the learners do not have a good understanding it could be detrimental to further mathematics learning.

1.1 Statement of the problem

The majority of learners are unable to retain an adequate amount of conceptual and procedural knowledge to achieve the minimum pass requirements for mathematics in grade 6 although the systemic results have been steadily improving throughout the past five years (from 26.5% in 2015 to 33.5% in 2018) (A. Hendricks, personal communication, March 18, 2019). A pass rate of 33.5% is still a cause of concern. This implies that 66.5% of learners were still unable to recall the processes or understand the mathematical concepts prescribed in the curriculum to achieve the minimum requirements.

The learner's inability to recall and retain mathematics concepts can more than likely be credited to the mass teaching practice approach that the South African Curriculum [called the Curriculum Assessment Policy Statement (CAPS)] follows (Kang, 2016). Mass practice can be defined as a teaching practice focused on a particular topic that requires a certain set of knowledge and skills, once these specific skills are considered learned, another topic is then introduced to be learned. Once the new topic is introduced the previous topics (as well as their skills) are neglected (Kang, 2016) (see figure 1).

Term 1	
Topic	**Time**
Mental Mathematics (10 minutes daily)	**8 hours**
Whole numbers: counting, ordering, comparing, representing and place value (6-digit numbers)	**2 hours**
Number sentences	**3 hours**
Whole numbers: addition and subtraction (5-digit numbers)	**7 hours**
Common fractions	**10 hours**
Time	**4 hours**
Properties of 2-D shapes	**8 hours**

Figure 1: An excerpt of the Intermediate phase time allocation per topic, in term 1 for grade 6. (Curriculum Assessment Policy Statement, page 212)

Figure 1 displays the time allocation as prescribed by CAPS for each topic. In my opinion, figure 1 is an indication that mass practice is the teaching method prescribed by CAPS, as each topic has a precise time frame allocated to the teaching and learning of a topic, thereafter, another topic will be learned. According to Rohrer, Dedrick, & Stershic (2015) mass practice allows learners to be able to predict the type of problems they will be asked to solve, as well as the method to solve them. For this reason, learners are not required to use their knowledge to identify which solution would fit their problem, but instead they predict what is being asked. Hence, it could be said that mass practice reduces the difficulty of problems (Rohrer, Dedrick & Stershic, 2015). This result in learners not being able to differentiate between problem types in tests and examinations. Consequently, learners would not develop the requisite conceptual understanding.

Mass practise allows for learners' acquired knowledge to be forgotten or fragmented once new topics are introduced (Rohrer and Taylor, 2010). Authors Hu, Liu, Chen, Liu, Yu, Deng, and Hosaka, S. (2013) contend (based on the arguments of (Ebbinghaus, 1885) that all people forget at a certain rate. They also state the only way to retard the forgetting process is to regularly review or practice the previously learned material. Regular practice is in direct contrast to mass practice. Therefore, it is my opinion, that the one way to enhance learner retention of their mathematical knowledge is to use an alternate teaching method called interleaved, distributed practice.

Interleaved, distributed practice is the blend of the two individual teaching strategies called interleaved practice and distributed practice. Interleaved practice is defined as the act of blending dissimilar skills to ensure that problems of a similar nature do not occur consecutively (Huges and Lee, 2019). Distributed practice is defined as the act of arranging practice sessions of skills, over a long period of time. In these sessions, the skills are practiced in smaller quantities than in the initial lessons (Huges and Lee, 2019). One method to improve conceptual understanding is to practise regularly over an extended period (Kang, 2016).

1.2 Purpose of the study

My goal for this study was to determine if a teaching strategy based on interleaved distributed practice would enhance the conceptual and procedural knowledge of participants to the extent that they will exceed the minimum requirements of the curriculum. In other words, I aimed to explore alternative teaching methods, other than those prescribed by the curriculum and other learning and teaching strategies. The idea of the intervention was to contribute to learners' conceptual and procedural understanding of mathematics and in the process to develop an understanding of how best to incorporate this into the curriculum prescribed by the department of education in South Africa. The intention was not only to enhance the knowledge retention and skills of the study participants' but in the process contribute to helping improve South Africa's quality of mathematics education.

1.3 Research Questions

1. How will distributed interleaved practice influence study participants' retention of mathematics knowledge?
2. How will distributed interleaved practice affect study participants' conceptual and procedural knowledge?

1.4 Conclusion

This chapter introduced the background of the study, stated the problem and purpose of the research, and ultimately stated the research questions that the study was be based on. A review of the literature to support the investigation will be presented in the next chapter.

Chapter 2: Literature Review

2.1 Introduction

Following the introduction of the study in chapter 1, this section describes the context of interleaved distributed practice as the subject of analysis in this thesis. It draws on existing literature to define the technique and to describe the significance and fundamental nature of this practice within mathematics education.

2.2 Interleaved distributed practice defined

The interleaved distributed practice differs from other learning practices, such as mass practice, as it is derived from kinesthetic research. Initially interleaving was used to examine and measure the effects it had on motor skills (Taylor and Rohrer, 2010; Hefferman, Hefferman and Ostow, 2015; Kang, 2016). Since then, the interleaved practice has been used by multiple researchers as an alternative to mass practice in mathematics education research (Roher and Taylor, 2010; Derick, et al., 2015; Hughes & Lee, 2019, Heffernan et al., Kang, 2016).

Interleaved distributed practice is derived from two different but inseparable learning techniques namely, interleaved practice and distributed practice (Huges and Lee, 2019).

Distributed practice is a learning technique that distributes, spaces or schedules learning over a long period of time. Findings of research have shown that distributed practice enhances long-term memory retention and, the application and transfer of skills (Dunlosky, Rawson, Marsh, Nathan, Willingham, 2013, p.35). Carey (2014) claims that distributed practice can double the amount of information retained by learners. I am not aware however of any study that investigated retention in terms of the prevalent mathematical knowledge types namely procedural and conceptual knowledge.

Interleaved practice is a teaching strategy where learners are presented with a mix of problem types, where consecutive problems are of a different type (Carey, 2014, p163). This means instead of having ten addition problems in one activity another topic would be interleaved within that same activity. Interleaving can assist in developing learners' ability to differentiate between different kinds of problems and to choose the appropriate strategies to solve these problems (Dunlosky et al.,2013; Rohrer and Taylor, 2010; Derick, Stershic and Roher, 2015; Hughes and Lee, 2019, Heffernan, Heffernan and Ostow, 2015; Kang, 2016).

Interleaved and distributive practices are suitable to be used in unison. This is because when tasks are interleaved problems are also separated or distributed over time (Carey, 2014). This statement presented a complication for researchers since distributed practice, on its own, as well as interleaved practice on its own has been shown to enhance test performance (Roher and Taylor, 2010; Derick et al., 2010; Hughes and Lee, 2019; Heffernan et al., 2015; Kang. 2016). Roher and Taylor (2010) contend that although previous studies did not account for this complication, it does not render their findings obsolete. They suggest that these studies simply be renamed to; "*interleaved distributed practice*". Furthermore, they argue that in the process of interleaving, distributing also takes place. Therefore, to measure interleaving independently they controlled the spacing. They did this by allocating equal amounts of spacing to the interleaved practice as they did the massed practice. The aforementioned act makes the Rohrer and Taylor (2010) study the first, to my knowledge, that measured interleaving while controlling for spacing. Other studies (Derick et al., 2010; Lee and Hughes, 2019; Heffernan et al.,2015; Kang. 2019), although aware of the confounding, combined the aforementioned teaching methods. The distributing effect within interleaving is one of the foremost reasons interleaving is seen to be effective (Pan, Tajran, Lovelett, Osuna, & Rickard, 2019)

Some studies examined the effects of interleaved practice on mathematics learning. Roher and Taylor (2010) presented a study on learners' ability to identify, the faces, edges, vertices, and angles of three-dimensional prisms. Their findings showed that interleaving increased the participant's performance. Derick et al., (2015) noted positive results for their study of interleaving learners' performance on linear equations and graphs. Heffernan et al., (2015) also found results in favour of interleaved distributed practice when they tested the skills involved in problem-solving of complementary, supplementary angles, the surface area of a pyramid, and probability of compound events without replacement. It is important to note that in all the aforementioned studies there was no difference in the teaching and learning conditions between massed practice and the interleaved practice group. The only difference was how these practice sessions had been organized.

For some researchers, the success of the interleaved practice can be credited to the aforementioned distributing effect as well as the Discriminative Contrast effect (Heffernan et al., 2015; Pan et al., 2018). The discriminative contrast effect refers to the contrast of the dissimilar tasks during the interleaving process as each lesson is followed by sets of problems drawn from many previous topics, so no two similar problems occur consecutively, thereby allowing learners to find solutions based on the problem itself (Heffernan et al., 2015; Pan et al., 2018). Therefore, it can be said that interleaving is based on problem type identification (Heffernan et al., 2015). According to Derick et al., (2015), this is an important aspect of mathematics as there are two significant parts to problem-solving. Firstly, identifying the problem type, then choosing an appropriate solution method, and finally implementing the chosen method. The interleaving practice, therefore, allow learners to choose the strategy for problem-solving based on each individual problem as opposed to massed practice where learners would have an idea of strategies to implement before even reading the problems (Derick et al., 2015).

Interleaved practice has been criticized as being difficult to implement in the classroom (Roher and Taylor, 2010; Herffernen et al., 2015). These authors argue that although the interleaved distributed practice yielded better test results than the mass practice results during formal assessments interleaving impaired the learners' performances during lessons. These authors called this phenomenon "desirable difficulty". This term was first established by Bjork (1994). Kang (2016) argues that interleaved practice gives the illusion of being more difficult in practice as the learners do not have the benefit of the repeated practice that is part of mass practice. Pan et al., (2018) however argue that increased effort during practice enhances long-term retention. Pan et al., (2018) and Dunlosky et al., (2013) suggest that massed practice used early on to strengthen skills could benefit the interleaving practice. While Kang (2016) contends that a hybrid approach can be beneficial, with new learning occurring via massed practice and interleaving used in a practice or consolidation phase.

Another disparity in the research was deciding on which cognitive area to interleave, be it the skill, the task type, or the representation (Heffernen et al., 2015). In this research, it was the intention to interleave and distribute presented problems. We believe that it would be a very difficult exercise to separate interleaved and distributed practice in a study done in authentic educational settings. This is since in the process of interleaving problems are also distributed. Therefore, the intention is to investigate interleaved and distributed practice as one phenomenon.

The literature reviewed indicates that interleaved and distributed practice helps to improve learner test performance. Knowledge retention in the reviewed literature is described in a general sense. None of the studies reviewed investigated the effects these practice types have on specific knowledge types that are important in mathematics learning namely conceptual and procedural knowledge as well as procedural flexibility.

2.3 Conclusion of chapter 2

Interleaved distributed practice apart from its difficulty to implement could be useful at all levels of mathematics. The potential contributions the intervention could promote in efficacy, skill, and knowledge by providing learners with the opportunity to learn how to choose a sufficient strategy in their mathematics problems would promote a desirable difficulty and thus promote learning. The various studies considered in this chapter bring forth a conclusion that there is a value in pursuing a study based on interleaved distributed practice. The next chapter (Chapter 3) will serve as the theoretical basis in pursuing a study based on interleaved distributed practice.

Chapter 3 Theoretical Framework

3.1 Introduction

This chapter explains the theoretical framework of this study. The theoretical framework will describe and explain the preexisting educational theories assumptions the study will be rooted in. This study explains how an interleaved distributed practice will improve conceptual and procedural knowledge by enabling procedural flexibility.

3.2 Procedural and Conceptual Knowledge

Proficiency in mathematics is dependent on a combination of three types of knowledge systems working interdependently. They are called, conceptual knowledge, procedural knowledge according to Rittle-Johnson & Schneider (2015) as well as procedural flexibility (Schneider, Rittle-Johnson, & Star, 2011).

Conceptual knowledge can be defined as, knowing more than just individual facts and procedures. It is having the ability to connect and adapt this knowledge in order to transfer it to multiple contexts and unknown situations (Rittle-Johnson & Schneider, 2015). Whereas, procedural knowledge can be defined as the performance and memorization of algorithms and procedures when solving problems, in many instances without making deeper connections with contexts or problems (Rittle-Johnson & Schneider, 2015).

Procedural flexibility can be defined learners having knowledge of more than one solution to problems and have the ability to adapt and apply these solutions in various contexts (Schneider, Rittle-Johnson, & Star, 2011). Procedural flexibility has two characteristics; the first being the ability to solve problems in multiple ways and the second, being the ability to choose the most fitting solution to the problem in a specific context (Schneider, Rittle-Johnson, & Star, 2011). Both conceptual and procedural knowledge aid in developing procedural flexibility.

3.3 Retrieval Effort Hypothesis

Interleaved distributed practice allows presented problems to also be distributed over time. The distribution causes learners to revise prior content in order to constantly have an active discourse with their prior knowledge (Bruner, 1966). This idea is strengthened by the theory of the *Retrieval Effort Hypothesis* (Pyc and Rawson, 2009). According to Pyc and Rawson, (2009) the Retrieval effort technique is the process of finding solutions to problems using knowledge that learners have previously learned. When attempting this process learners recall previously learned knowledge. This process helps strengthen their memory and improves their ability to apply the knowledge. Retrieving this knowledge reactivates the ideas that were involved in processing the events that occurred when the ideas were first learnt (Frank and Macnamara, 2017). This would allow the learner to think about the information actively and purposefully they are trying to learn and bring it to the forefront of their minds. Pyc and Rawson (2009) also, state that although any retrieval practice is successful, difficult retrieval is better to help learners retain long-term memory. This idea advances the "desirable difficulty" of interleaved distributed practice as explained by Bjork (1994), Kang (2016) and Pan et al., (2018) in Chapter 2. Frank and Macnamara (2017) agrees and states that retrieval strategies improve efficiency above and beyond what is gained by practicing the procedure alone. They also state that in order for the retrieval process to be valuable in increasing retention, three phases need to occur. Within each successive stage errors made and time taken to complete these tasks decreases. These are called the learning phase; difficulty phase and critical phase. Within the learning phase information is learnt through algorithms and practiced then stored away in the short term memory. In the difficulty phase the problems presented to the learners in a mixed and spaced fashion with unrelated topics to follow one another. In this difficulty phase learners are required to look at their problems individually rather than algorithmically. In the critical phase the difficulty level stays the same as within the difficulty phase and learners would perform better. This too suggests that "desirable difficulty" is a necessary step within interleaved practice in order to enhance memory.

3.4 Cognitive Load Theory within the Distributed practice domain

The Cognitive Load Theory (CLT) is an instructional theory that explains how the human mind obtains original knowledge and how original knowledge transcends the working memory and grows into long term memory (Chen, Pass & Sweller, 2021). New information is learned from others such as teachers or peers, gathered from worldly experiences or self-generated unsystematically. The information is then processed, used as working memory -for a limited time period- and will be verified within various contexts such as problem-solving. Based on these experiences fragments of this information is then processed and stored for an indeterminate amount of time in the long term memory. Long term memory can be used as schema to recall when needed, to solve more intricate problems (Chen, Pass & Sweller, 2021). Chen, Pass & Sweller, (2021) states that working memory having a limited capacity and disrupts the acquiring of knowledge process. This means that when the working memory is overloaded either by heavy cognitive effort or if information shares similar cognitive components. The Working memory develops a limit and the process between working memory and long term memory breaks down. CLT argues that this break down can be amended by the "rest –from- deliberate learning" that is presented within the space within the interleaved distributed practice. Chen and Kalyuga (2020) states suspending assessments would allow working memory resources to amend whereas immediate assessment would insert cognitive load to a working memory that has already been exhausted by the learning activities that came before. This means that time is used as the key factor that impacts working memory reduction and restoration. The rest-from-deliberate-learning to aid with working memory favours interleaved distributed practice as it allows the learner to take a break from the topic being learnt through the "spacing effect". When the learner then encounters the topic again through the interleaving effect it allows the learner to retrieve the information to solve the problems, which allows retrieval. In our teaching strategy rest-from-deliberate learning was done by distributive practice. That is learning of a concept is done over multiple sessions. In other words there are 'rest' or time in between learning sessions.

3.5 The discriminative-contrast hypothesis within the interleaved domain

The discriminative-contrast hypothesis is founded on the assumption that interleaved distributed practice is presented through necessary difficulty by the discrimination and comparison process (Nemeth, Werker, Arend, & Lipowsky, 2021). During interleaved distributed practice, problems of different types are mixed up during the intervention process. This act aids learners' ability to develop the ability to differentiate between dissimilar problem types and to make cognitive links between problem types and an appropriate solution procedure (Derick et al., 2015; Carvalho and Goldstone, 2015). Therefore, it can be presumed that interleaved distributed practice promotes the different dimensions of knowledge, such as how to apply various procedures within different contexts and also when and why to use these procedures (Nemeth et al., 2021). Thus, it may be argued that interleaved distributed practice supports the development of procedural flexibility. The goal of mathematics teaching is to develop conceptual and procedural knowledge that can be retained in the long term and is flexible in various contexts (Suzuki, Nakata and Dekeyser, 2019).

Chapter 4: Research Methodology

4.1 Introduction

This chapter presents the research design utilized to determine the effects an interleaved distributed practice teaching approach will have on levels and retention of conceptual and procedural knowledge and procedural flexibility. It includes a detailed explanation of the research methodology, the data collecting methods, and the data analysis tools employed.

4.2 Methodology

A quasi-experimental research design was utilized in the study. The quasi-experimental research design is an empirical intervention used to estimate the impact of an intervention on a target population that is grouped together non-randomly (Chiang, Jhangiani & Price 2015).

According to Chiang et al., (2015) a quasi-experimental design resembles true quantitative research since it observes and reports on behaviour before and after specific intervention(s). It is however different from a true experimental design in the following way:

> "*Because the independent variable is manipulated before the dependent variable is measured, quasi-experimental research eliminates the directionality problem. But because participants are not randomly assigned—making it likely that there are other differences between conditions*" (Chiang et al., 2015, p1)

This means that the sample groups that were studied were organized in a pre-existing way and therefore the researcher had no control over these conditions.

The quasi- experimental method was ideal for this research since research participants were not randomly assigned to groups, but pre-existing groups (class groups) were utilized. The groups that were studied were organized in specific classes, over which the researcher had no control. However, it always was the intention to do the research in an authentic educational setting and so any disturbance of the class groups (random assignment) was avoided.

A pre-and post-test control and experimental group design were followed in the study. Data is participant test scores. Data was collected at specified points before and after the intervention. The intervention (interleaved distributed practice) is the independent variable and the test scores of participants are the dependent variable. The independent variable was manipulated (this will be discussed in more detail later) and was implemented after participants had been exposed to teaching on specified topics. Three class groups from the researcher's school were employed as study participants. For certain parts of the study, one group was utilized as the control group while the other group was the experimental group in other parts the experimental and control groups were inverted (this will also be explained in more detail later) and the third group was used as a baseline group where no intervention was used. The control, experimental, and baseline group scores were compared and subjected to statistical analysis.

4.3 Sampling

A nonequivalent group (Campbell & Riecken, 1968) design was utilized in this study. The quasi-experimental research design allows for the sampling method for the study to be organized in a non-random and specific way. Therefore, because of this non-random grouping the groups that were studied whether it be the control group or the experimental groups the groups were dissimilar (Chiang et al., 2015). The term 'nonequivalent groups design' is rooted in the dissimilar design of the target population. The groups to be studied are organized in classes by the school administration staff in a specific way.

The sample size was 116 grade 6 learners. The number of learners in the first class (called group A) was 39 and the number of learners in the second class (called group B) was 39 and the number of learners in the third class (called group C) was 38. The gender demographics in the groups was as follows: Group A has 14 boys and 25 girls group B has 23 boys and 16 girls Group C was 19 boys and 19 girls. The breakdown of demographics is shown in table 1 below.

	Boys	Girls	Average age
Group A	14	25	12
Group B	23	16	12
Group C	19	19	12

Table 1: Demographics

4.4 Research instruments

The tool used to conduct this research that will be used is a pre and post test design of three groups of learners.

4.4.1 Pre-tests and Post tests

In order to determine the effect of the *interleaved distributed practice*, learners were presented with tests prepared by the researcher. The same test was written by both the control group, experimental group, and baseline group. These test questions were based on the topic of fractions. This is a topic that is usually done in the second term. Questions of varying difficulty levels were posed. Tests were set by the researcher which was moderated by the supervisor of the study. Marking was done by the researcher and moderated by the supervisor. This was done in order to deal with possible researcher bias. The cognitive demand (in terms of knowledge and reasoning level) of each question was determined by means of taxonomy (see table 2 below)

The day after the teaching of a subtopic was completed participants were presented with a test. This test was considered to be a pre-test. Following the pre-test, the intervention in the form of interleaved and distributed practice was implemented. Approximately 3 to 4 weeks after the pre-test participants were presented with the post-test.

See Appendix A for the research instruments.

4.4.2 Taxonomy Table

<table>
<tr><th rowspan="4">The knowledge Dimension (KD)</th><th colspan="5">The cognitive process dimension (CPD)</th></tr>
<tr><td colspan="3">1. Imitative reasoning (IM)</td><td colspan="2">2. Creative mathematically founded reasoning</td></tr>
<tr><td rowspan="2">(a.) Memorised reasoning (MR)</td><td colspan="2">(b.) Algorithmic Reasoning (AR)</td><td rowspan="2">(a.) Local creative reasoning</td><td rowspan="2">(b.) Global creative reasoning</td></tr>
<tr><td>(i) familiar AR</td><td>(ii) delimiting AR</td></tr>
<tr><td>A.) Factual knowledge</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>B.) Procedural knowledge</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>C.) Flexible procedural knowledge</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>D.) Conceptual knowledge</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Table 2: The revised taxonomy table (May, 2021, p.6).

4.5 Teaching strategy for data collection

As mentioned previously the research utilized three classes. These classes are labelled Group A, Group B, and Group C. Group A and B served as a control group and an experimental group, and the third (Group C) was used as a baseline group. The baseline group would experience no intervention. The experimental and control groups received lessons on the topic in the same manner by the same teacher.

The topic of Fractions was divided into 4 subtopics. For the first subtopic group, A was the experimental group, and group B was the control group. For subtopic 2 group A was the control group and group B the experimental group and so on. Immediately following the completion of the teaching of each subtopic a test was written. This test was considered a pre-test for the subtopic the pre-test served as a means to benchmark learner understanding. Next, both groups continued with the succeeding topics as prescribed in the curriculum. During these subsequent lessons, the experimental group continued with interleaved and distributed practice of the subtopic daily and the control group continued without this for the specific topic. This process took place from 18 April to -24 June 2022. Following this time period, both groups wrote a post-test to measure both groups' development. This process was followed for all the subtopics in fractions and decimals.

To ensure that neither group was completely disadvantaged, both groups were at some stage exposed to the interleaved distributed practice. (See diagram below)

Topics to be taught	Group A	Group B	Group C
Topic 1: Addition and Subtraction of proper fractions	Experimental Group	Control Group	No Intervention
Topic 2: Adding and subtracting mixed numbers	Control Group	Experimental Group	No Intervention
Topic 3: Fractions of an amount	Experimental Group	Control Group	No Intervention
Topic 4: Percentages of whole numbers	Control Group	Experimental Group	No Intervention

Table 3: The teaching strategy for experimental and control groups respectively

4.6 Research implementation

The investigation utilized 30-minute periods over a 10 week period. During the intervention, the learner's fraction knowledge was periodically assessed by observation during class time by checking the learners' writing books and giving feedback verbally by doing corrections. In each blocked lesson, the participants were asked to engage in the lesson to answer questions, discuss the rules, and complete the activity after the lesson was completed. In every fourth period, there would be an interleaving period where the previous topics were reviewed.

4.7 Conclusion of Chapter 4

This chapter explained the research methodology used by the researcher in this study. A quasi-experimental research design was implemented using three grade 6 classes as participants. The classes were studied throughout a term as a control and experimental group. A pre-test was presented after the learning of subtopics and a post-test was presented after an interleaved distributed practice had been implemented.

Chapter 5: Statistical analysis

5.1 Introduction

Pre and post-test test results were compared in order to determine whether the learners managed to gain a procedural flexibility and conceptual understanding. The IBM SPSS (version 28) software was used to do the statistical analysis. For both the pre-and post-tests individual questions were categorized using a revised taxonomy table (May, 2022, p.6) (see table 2). We added the scores for the variables A1a, B2a, B1bi and C1bi. This was done separately for class groups A, B and C. These scores were deemed to represent a measure of the competency procedural fluency. The variable is represented in the statistical analysis with the word '*skill*'.

We also added the scores for the variables D1bi, D1a and D2a. The scored were added separately for class groups A, B and C. These scores were deemed to represent a measure of the competency conceptual understanding. The variable is represented in the statistical analysis with the word *'conc'*.

5.2 Descriptive statistics

Before proceeding with the statistical analysis of the data we explored it by means of descriptive statistics. This was done in order to check for possible violations of underlying assumptions in statistical tests. We determined the following descriptive statistics for the data: mean, standard deviation, range, skewness and kurtosis. Histograms were utilized in order to present a visual display of the distribution of the data. Descriptive statistics were done for the following variables. Variables for the pre-test is: control skill pre; control conc pre; exp skill pre; exp conc pre; whereas variables for the post-test is: control skill post; control conc post; exp skill post and exp conc post. Different variables represent class groups A, B and C.

The following holds for skewness and kurtosis. Skewness and kurtosis provide information concerning the distribution of scores on continuous variables. The skewness value offers an indication of the symmetry of the distribution whereas the kurtosis provides information about the 'peakedness' of the distribution. Positive skewness values indicate scores clustered to the left at low values whereas negative skewness values indicate a clustering of scores at the high values. Positive kurtosis values indicate that the distribution is clustered in the centre (peaked) with long thin tails whereas negative kurtosis values indicate a distribution that is relatively flat. If the distribution is perfectly normal, both skewness and kurtosis will have a value of 0. Histograms are used to show this visually.

5.2.1 Descriptive statistics pre-test

It is communal practice that when undertaking a statistical analysis one explores the records by means of descriptive statistics and graphs as a means to thoroughly describe the data. This process is essential to comprehend the data as well as to check for any contraventions of underlying assumptions in the statistical tests. These checks take place occur as one needs to ensure that the data is normally distributed and if any outliers exist, these checks are necessary as they may influence correlation coefficients. The descriptive statistics of the mean, standard deviation, range, skewness and kurtosis were obtained from the experimental groups A and B SKILLPRE, SKILLPOST, CONCPRE, CONCPOST as well as the control group C SKILLPRE, SKILLPOST, CONCPRE and CONCPOST.

5.2.1.1 Control Group A PRE

Descriptive Statistics control group A Pre test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CONTROL A SKILL PRE	42	42.1	.0	42.1	16.040	1.844	11.9512	142.83	.535	.365	-.272	.717
CONTROL A CONC PRE	42	60	0	60	18.33	2.682	17.379	302.033	.971	.365	.324	.717
Valid N (listwise)	42											

Table 5.1 Descriptive statistics for control Group A pre-test

5.2.1.2 Histogram control group skill A pre test

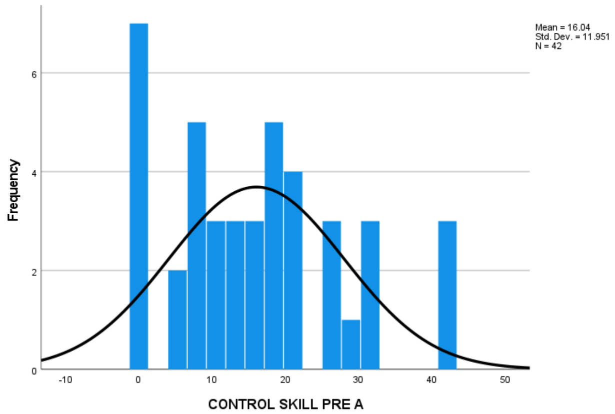


Figure 5.1: Histogram CONTROL A SKILLPRE

Histogram control group Conc A Pre test

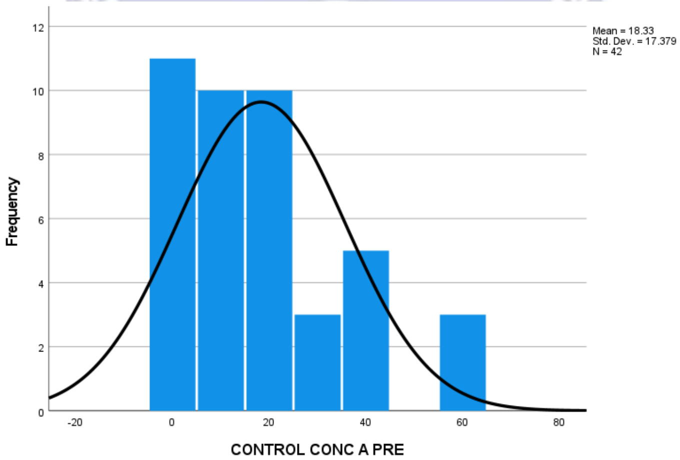


Figure 5.2: Histogram CONTROL A CONCPRE

5.2.1.3 Control group B Pre-test

Descriptive Statistics Control group B Pre-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CONTROL B SKILL PRE	43	53.6	.0	53.6	19.352	2.1630	14.1840	201.186	.889	.361	.245	.709

CONTROL B CONC PRE	43	42.9	.0	42.9	16.113	1.9901	13.0497	170.294	.363	.361	-.660	.709
Valid N (listwise)	43											

Table 5.2 Descriptive statistics for control Group B pre-test

5.2.1.4 Histogram Control Skill B Pre

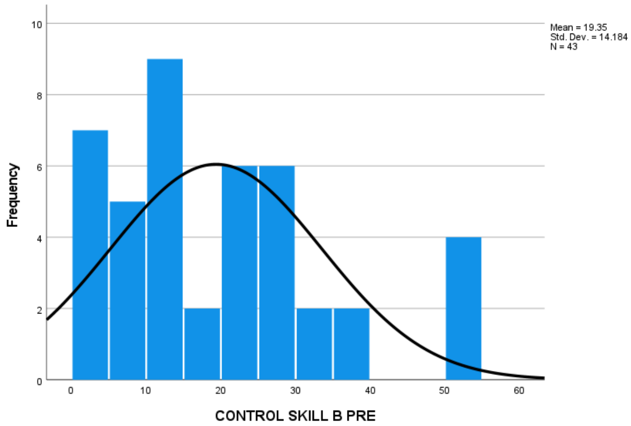


Figure 5.3: Histogram CONTROL B SKILLPRE

5.2.1.5 Histogram Control Conc B Pre

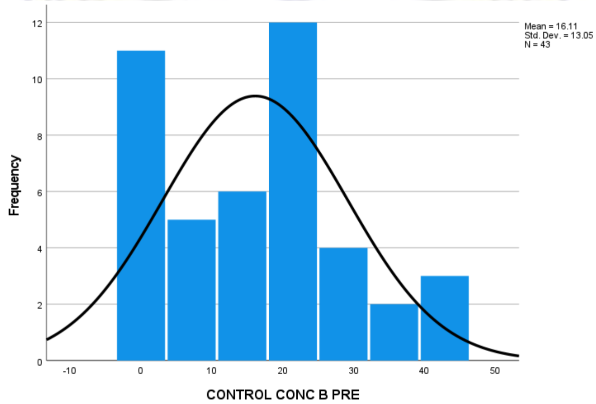


Figure 5.4: Histogram CONTROL B CONCPRE

5.2.1.6 Experimental group A Pre test

Descriptive Statistics of experimental group A Pre-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
EXP A SKILL PRE	43	100.0	.0	100.0	25.332	3.3082	21.6936	470.611	1.228	.361	2.173	.709
EXP A CONC PRE	43	100.0	.0	100.0	20.598	2.9613	19.4187	377.088	1.747	.361	5.398	.709
Valid N (listwise)	43											

Table 5.3 Descriptive Statistics of experimental group A Pre-test

5.2.1.7 Histogram EXP SKILL A

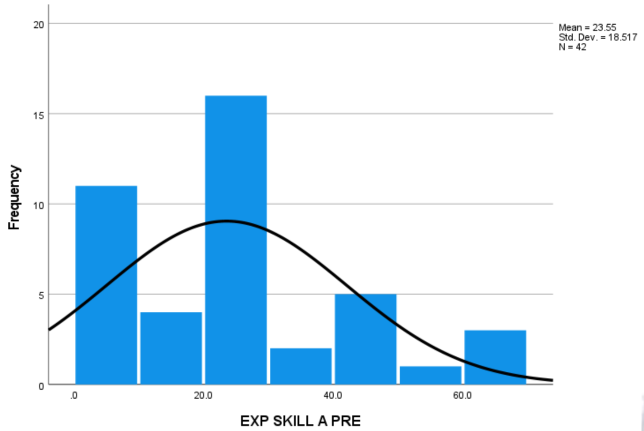


Figure 5.5 Histogram EXPERIMENTAL A SKILLPRE

5.2.1.8 Histogram EXP CONC A

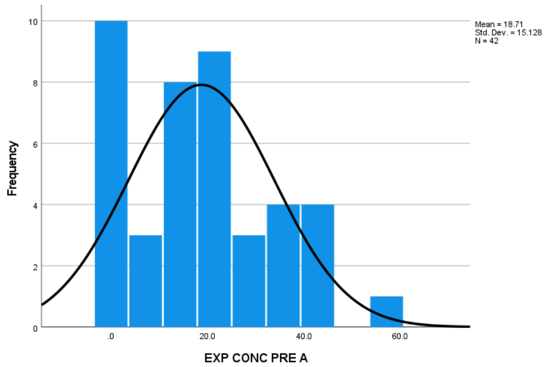


Figure 5.6 Histogram EXPERIMENAL A CONCPRE

5.2.1.9 Experimental group B Pre test

Descriptive Statistics of experimental group B Pre-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
EXP B SKILL PRE	43	39.5	.0	39.5	14.933	1.4489	9.5008	90.265	.589	.361	-.263	.709
EXP B CONC PRE	43	50	0	50	14.65	2.166	14.201	201.661	.676	.361	-.593	.709
Valid N (listwise)	43											

Table 5.4 Descriptive Statistics of experimental group B Pre-test

5.2.1.10 Histogram EXP SKILL B Pre

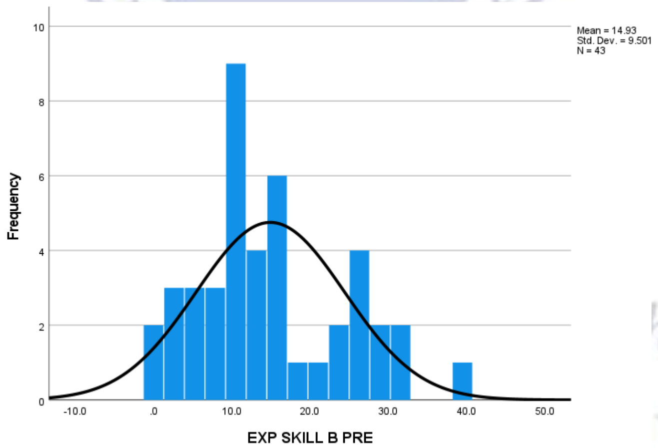


Figure 5.7 Histogram EXPERIMENAL B SKILLPRE

5.2.1.11 Histogram CONC B PRE

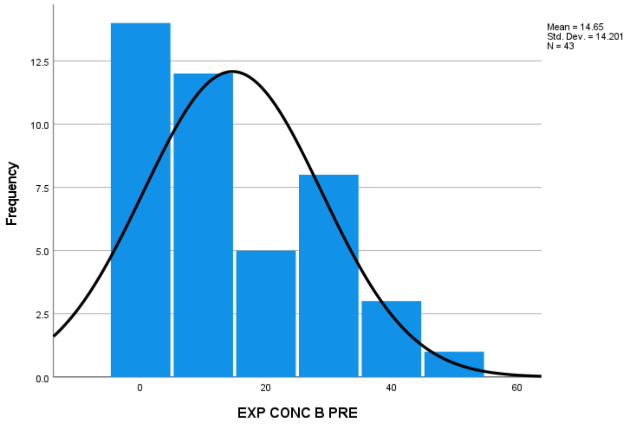


Figure 5.8 Histogram EXPERIMENTAL B CONCPRE

5.2.1.12 CONTROL GROUP C PRE-TEST

Descriptive Statistics of experimental group C Pre-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CONTROL C SKILL PRE	40	25.8	3.0	28.8	14.242	.9171	5.8006	33.646	.438	.374	-.290	.733
CONTROL C CONC PRE	40	29.2	4.2	33.3	16.667	1.1842	7.4893	56.090	.167	.374	-.615	.733
Valid N (listwise)	40											

Table 5.5 Descriptive Statistics of experimental group C Pre-test

5.2.1.13 HISTOGRAM SKILL CONTROL C PRE

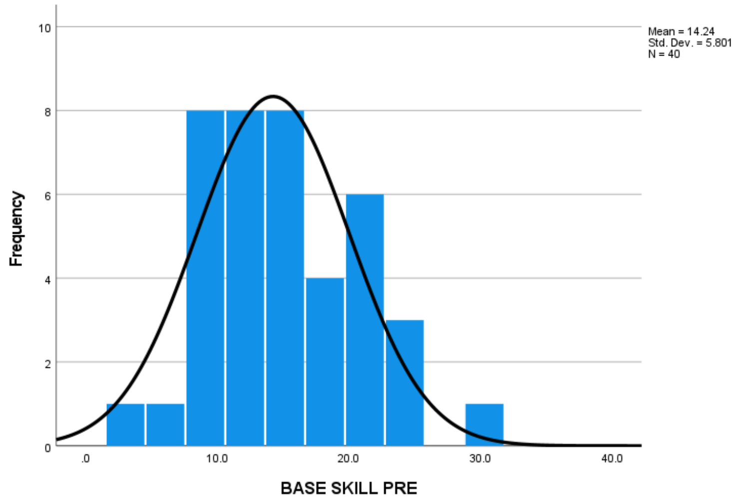


Figure 5.9 Histogram CONTROL C SKILL PRE

5.2.1.14 Histogram CONC control C PRE

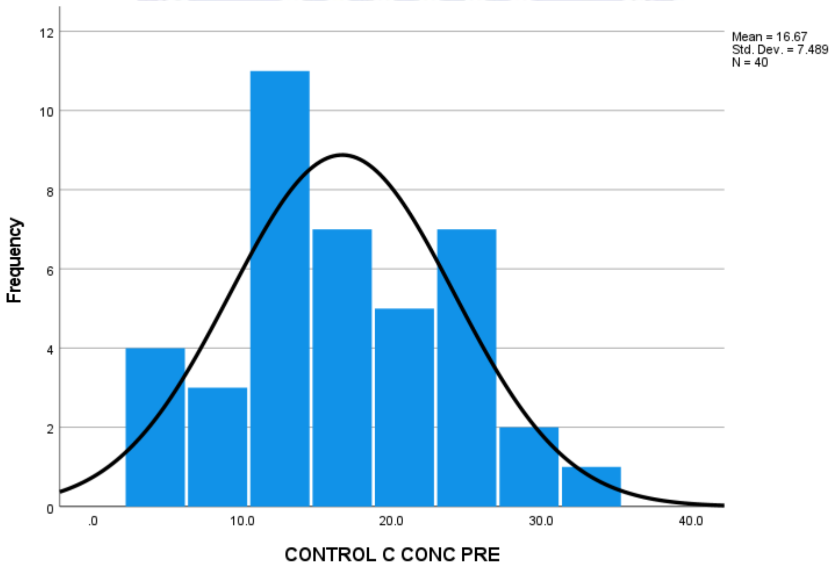


Figure 5.10 Histogram CONTROL C CONCPRE

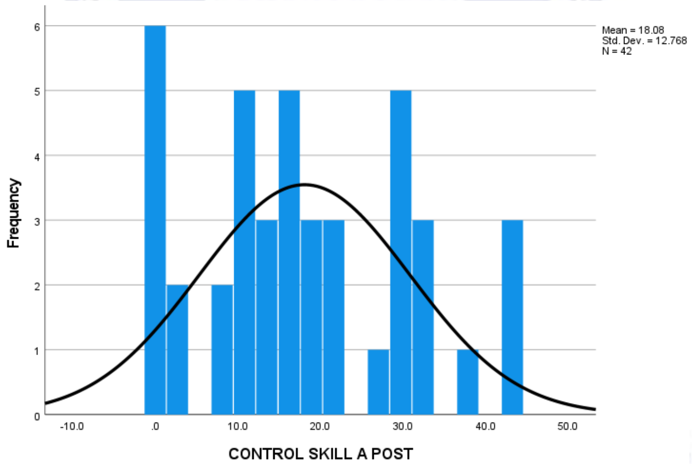
5.2.2 Descriptive statistics: Post-test

Descriptive Statistics control group A Post test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CONTROL A SKILL POST	42	43.2	.0	43.2	18.082	1.9701	12.7677	163.013	.322	.365	-.715	.717
CONTROL A CONC POST	42	40	0	40	17.14	2.163	14.018	196.516	.261	.365	-1.183	.717
Valid N (listwise)	42											

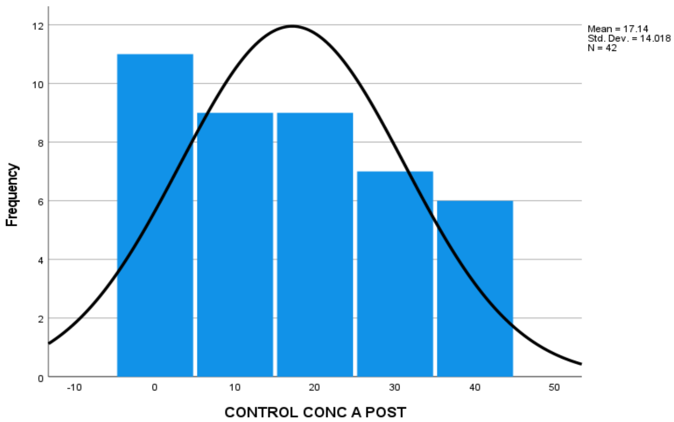
Table 5.5 Descriptive Statistics control group A Post test

5.2.2.1 Histogram Control SKILL A POST



Graph 5.11 Histogram CONTROL A SKILLPOST

5.2.2.2 Histogram Control A CONC POST



Graph 5.12 Histogram CONTROL A CONCPOST

5.2.2.3 CONTROL GROUP B POST-TEST

Descriptive Statistics control group B Post test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CONTROL B SKILL POST	43	53.1	.0	53.1	19.477	1.9972	13.0966	171.520	.926	.361	.377	.709
CONTROL B CONC POST	43	50	0	50	14.42	2.032	13.328	177.630	.637	.361	-.448	.709
Valid N (listwise)	43											

Table 5.6 Descriptive statistics Control group B post

5.2.2.4 Histogram Control group B skill post test

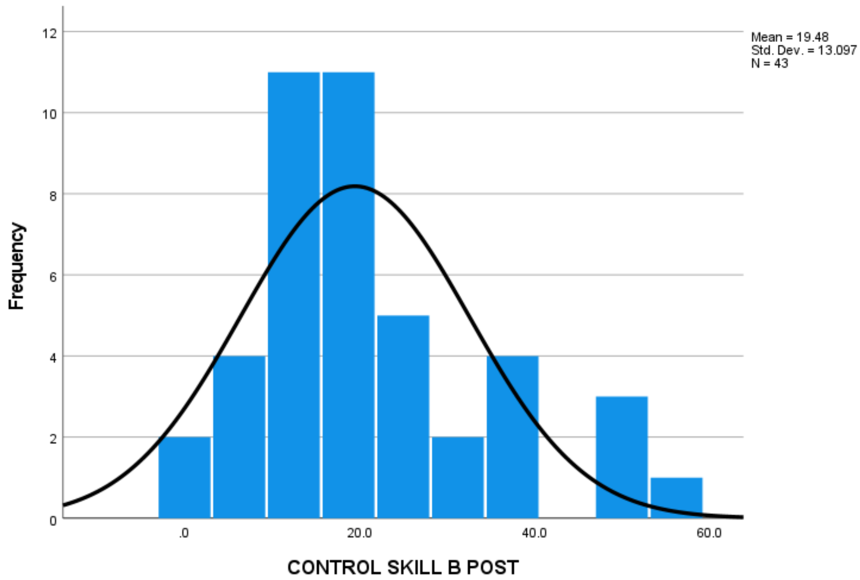


Figure 5.13 Histogram CONTROL SKILL B POST

5.2.2.5 Histogram CONTROL CONC B POST

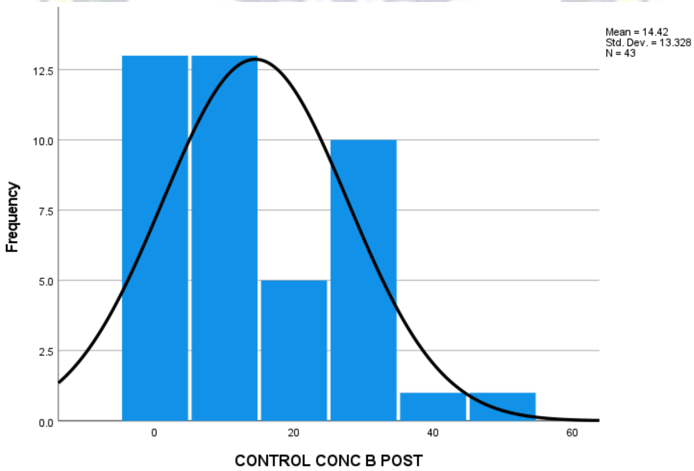


Figure 5.14 Histogram CONTROL B CONCPOST

5.2.2.6 Experimental group A Post-Test

Descriptive Statistics of experimental group A Post-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
EXP A SKILL POST	42	60.7	3.6	64.3	27.721	2.4246	15.7129	246.895	.501	.365	-.132	.717
EXP A CON POST	42	50.0	.0	50.0	21.429	2.1737	14.0870	198.443	.507	.365	-.602	.717
Valid N (listwise)	42											

Table 5.7 Descriptive Statistics of experimental group A Post-test

5.2.2.7 Histogram EXP SKILL A POST

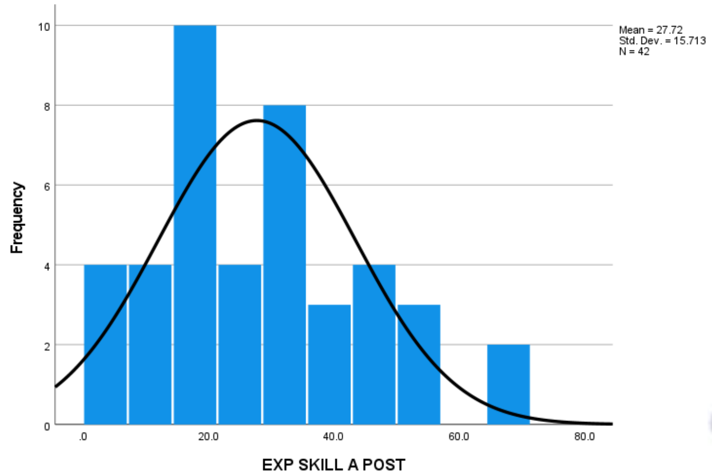


Figure 5.15 Histogram EXPERIMENTAL A SKILLPOST

5.2.2.8 Histogram EXP CONC A POST

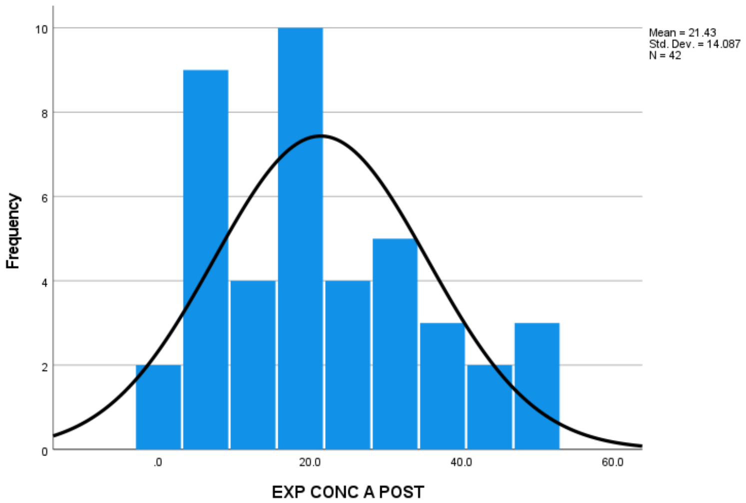


Figure 5.16 Histogram EXPERIMENTAL A CONCPOST

5.2.2.9 Experimental group B Post-Test

Descriptive Statistics of experimental group B Post-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
EXP B SKILL POST	43	37.8	5.4	43.2	21.245	1.5722	10.3098	106.293	.455	.361	-.541	.709
EXP B CONC POST	43	70	0	70	18.84	2.311	15.152	229.568	1.024	.361	1.524	.709
Valid N (listwise)	43											

Table 5.8 Descriptive Statistics of skill for experimental group B Post-test

5.2.2.10 Histogram EXP SKILL B POST

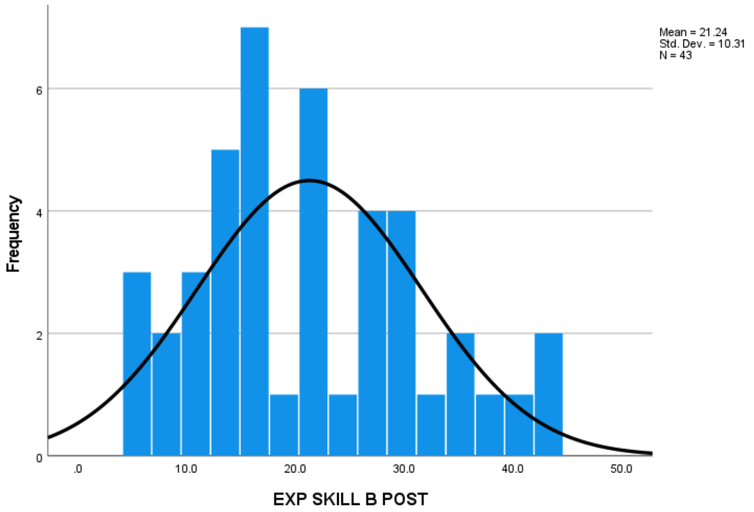


Figure 5.17 Histogram EXPERIMENTAL B SKILLPOST

5.2.2.11 Histogram EXP CONC B POST

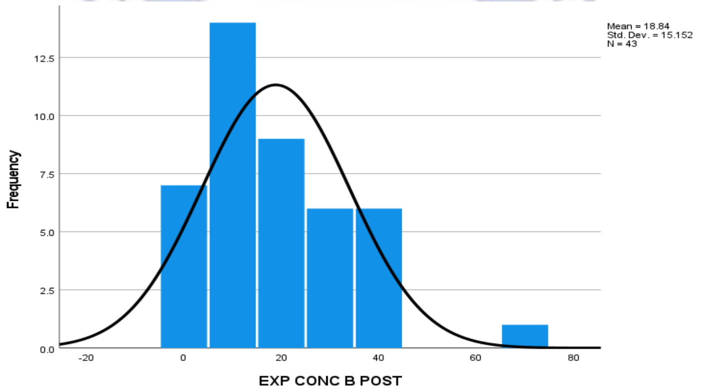


Figure 5.18 Histogram EXPERIMENTAL B CONCPOST

5.2.2.12 CONTROL GROUP C POST TEST

Descriptive Statistics of Control group C Post-test

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CONTROL C SKILL POST	40	21.7	2.9	24.6	13.514	.8438	5.3365	28.478	.305	.374	-.506	.733
CONTROL C CONC POST	40	35	0	35	14.00	1.293	8.181	66.923	.413	.374	.221	.733
Valid N (listwise)	40											

Table 5.9 Descriptive Statistics of control group C Post-test

5.2.2.13 Histogram skill Control C Post

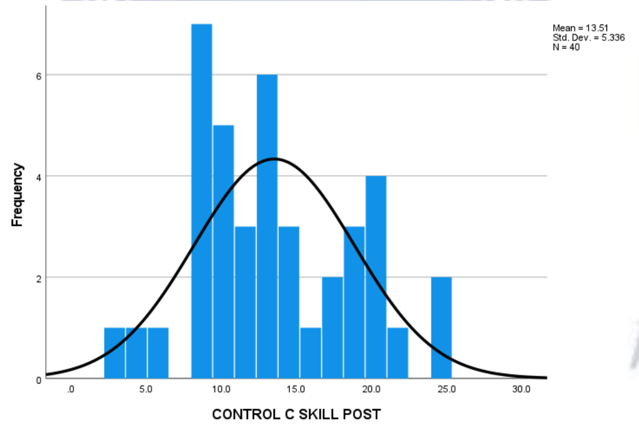


Figure 5.19 Histogram CONTROL C SKILLPOST

5.2.2.14 Histogram CONC control C post

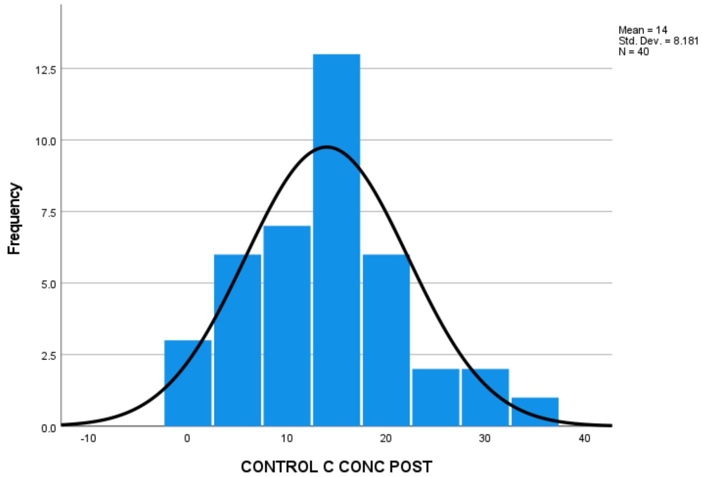


Figure 5.20 Histogram CONTROL C CONCPOST

The skewness of the value for the variables CONTOL A CONCPRE, CONTROL B SKILLPRE, EXPERIMENTAL A SKILLPRE, EXPERIMENTAL A CONCPRE. EXPERIMENTAL B CONCPRE, CONTROL B SKILLPOST are skewed positively which indicates that they are clustered at the low scores. The kurtosis for these variables are positive which indicates that scores are peaked around the mean of the data. The skewness and kurtosis taken together for the variables therefore indicate that most scores for these variables peak at the low scores.

The skewness of the values for the variables EXPERIMENTAL B CONCPRE, CONTROL B CONCPOST and EXPERIMENTAL A SKILLPOST are skewed positively which indicates that they are clustered at the low scores. The kurtosis for these variables are negative which indicates that the distribution of scores are flat and that most scores are in the extremes. Skewness and kurtosis taken together for these variables indicate that most scores for these variables are low.

A normal distribution has a bell-shaped curve, which indicates that the majority of the scores are in the middle with less frequencies towards the extreme. Such is the case for CONTROL A SKILLPRE, CONTROL B CONCPRE, EXPERIMENTAL B SKILLPRE, CONTROL C SKILL PRE, CONTROLL C CONCPRE, CONTROL A SKILLPOST, CONTROL A CONCPOST, EXPERIMENTAL A CONCPOST, and EXPERIMENTAL B SKILLPOST. The kurtosis for these variables are negative which is an indication that the distribution of scores are flat and that most scores are in the extreme. Skewness and kurtosis taken together for these variables indicate that most scores are clustered around the mean and the extremes spread outward from the mean.

A normal distribution has a bell-shaped curve, which indicates that the majority of the scores are in the middle with smaller frequencies towards the extreme. Such is the case for EXPERIMENTAL B CONCPOST and CONTROL C CONCPOST. The kurtosis for these variables are positive which indicate that the distribution is clustered at the centre. Skewness and kurtosis taken together for the variables indicate that most scores are clustered around the peaks are less within the extremes of the data set.

Therefore since the data do not violate the normality criteria too severely parametric statistics is preferred.

5.3 Statistical analysis

In the statistical analysis that follows we compared control groups pre- and post-test and experimental groups pre- and post-test (this was done separately for skill and conc scores). This was done to determine if there is a significant difference in mean scores. Paired-samples t-tests was used for this purpose.

We also compared mean scores of control and experimental groups post-test (this was done separately for skill and conc scores). This was done to determine if there is a significant difference between control and experimental groups post intervention. Independent samples t-tests was utilized for this purpose.

5.3.1 Paired-samples t-tests

For the paired-samples t-test the null hypothesis was that there is no significant difference after exposure to the teaching strategy. In other words, the mean difference of the pre- and post-test score for the population is zero i.e.

$$H_0: \mu_D = 0$$

The alternative hypothesis is that the intervention caused the post-test scores to be higher or lower than the pre-test scores. In other words, the mean difference is not zero:

$$H_1: \mu_D \neq 0$$

The level of significance is set at $\alpha = .05$ for a two-tailed test.

5.3.1.1 Paired Samples Statistics of Experimental group A pre-test and post-test

The output for the paired-samples test EXP A PRE VS EXP A POST is presented in the tables below:

Paired Samples Statistics of Experimental group A Pre-test to post-test

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	EXP SKILL POST	29.402	43	19.0397	2.9035
	EXP SKILL PRE	25.332	43	21.6936	3.3082
Pair 2	EXP CONC POST	23.256	43	18.3654	2.8007
	EXP CONC PRE	20.598	43	19.4187	2.9613

Table 5.10: Experimental group A Pre-test compared post-test

Paired Samples Statistics of Experimental group A pre-test and post-test

	Paired Differences					t	df	Significance	
				95% Confidence Interval of the Difference					
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1 EXP A SKILL POST - EXP A8 SKILL PRE	4.069	8.2324	1.2554	1.5362	6.6033	3.242	42	.001	.002
Pair 2 EXP A CON POST - EXP A8 CONC PRE	2.657	12.7473	1.9439	-1.2652	6.5809	1.367	42	.089	.179

Paired Effect Sizes of Experimental group A pre-test and post test

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Pair 1	EXP SKILL POST - EXP SKILL PRE	Cohen's d	8.2324	.494	.175	.809
		Hedges' correction	8.3069	.490	.173	.801
Pair 2	EXP CON POST - EXP CONC PRE	Cohen's d	12.7473	.208	-.095	.509
		Hedges' correction	12.8626	.207	-.094	.505

a. The denominator used in estimating the effect sizes.

Cohen's d uses the sample standard deviation of the mean difference.

Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Table 5.11 Paired sample statistics of experimental group A pre-test and post- test

Table 5.12 Paired Effect Sizes of Experimental group A pre-test and post test

A paired-samples t-test was conducted to evaluate the impact on students' scores of a teaching intervention based on interleaved and distributed practice. There is a statistically significant increase in scores for the variable EXP A skill from pre- test $(M = 25.33,\ SD = 19.04)$ to post-test $(M = 29.40, SD = 21.69), t(42) = 3.24, p = .002\ (two - tailed)$.The mean increase in scores for the skill variable is 4.07 with a 95% confidence interval ranging from1.54 to 6.60. The value for Cohen's d is 0.49 which is close to a medium effect size. The null hypothesis for the skill variable is rejected that is $H_1: \mu_D \neq 0$. We conclude that the teaching intervention contributed to an increase in post scores for the skill variable for experimental group A

There was not a statistically significant increase in scores for the variable EXP A conc from pre- test $(M = 20.6,\ SD = 18.37)$ to post-test $(M = 23.26, SD = 19.42), t(42) = 1.37, p = .179\ (two - tailed)$.The mean increase in scores for the conc variable was 2.66 with a 95% confidence interval ranging from -1.265 to 1.367. The increase for the conc variable however is not statistically significant. The null hypothesis is rejected but we conclude that the intervention did not significantly enhance conc post scores for experimental group A.

5.3.1.2 Paired Samples Statistics of Experimental group b pre-test and post-test

Output for the paired-samples test EXP B Pre vs. EXP B POST is presented in the tables below:

Paired Samples Statistics of experimental group B pre-test-post test

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SKILL EXP POST B	21.245	43	10.3098	1.5722
	SKILL EXP PRE B	14.933	43	9.5008	1.4489
Pair 2	CONC EXP POST B	18.84	43	15.152	2.311
	CONC EXP PRE B	14.65	43	14.201	2.166

Table 5.13 Paired sample statistics of experimental group B pre-test and post- test

<u>Paired Samples Test Experimental group B</u>

	Paired Differences					t	df	Significance	
				95% Confidence Interval of the Difference					
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1 SKILL EXP POST B - SKILL EXP PRE B	6.3118	4.1687	.6357	5.0289	7.5948	9.928	42	<.001	<.001
Pair 2 CONC EXP POST B - CONC EXP PRE B	4.186	5.448	.831	2.509	5.863	5.039	42	<.001	<.001

Table 5.14 Paired sample statistics of experimental group B pre-test and post- test

Paired Effect Sizes experimental group B Pre-test and post test

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	SKILL EXP POST B - SKILL EXP PRE B	Cohen's d	4.1687	1.514	1.070	1.950
		Hedges' correction	4.2064	1.501	1.060	1.932
Pair 2	CONC EXP POST B - CONC EXP PRE B	Cohen's d	5.448	.768	.424	1.106
		Hedges' correction	5.497	.762	.420	1.096

a. The denominator used in estimating the effect sizes.

Cohen's d uses the sample standard deviation of the mean difference.

Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Table 5.15 Paired Effect Sizes experimental group B Pre-test and post test

There was a statistically significant increase in scores for the variable EXP B skill from pre-test $(M = 14.93,\ SD = 9.50)$ to post-test$(M = 21.25, SD = 10.31), t(42) = 9.93, p <$ $.001\ (two - tailed)$. The mean increase in scores for the skill variable was 6.31 with a 95% confidence interval ranging from 5.03 to 7.60 . The Cohen's d statistic (1.514) indicated a large effect size. The null hypothesis for the skill variable is rejected that is$H_1: \mu_D \neq 0$. We conclude that the teaching intervention contributed to an increase in post scores for the skill variable for experimental group B.

There is a statistically significant increase in scores for the variable EXP B conc from pre- test $(M = 14.65,\ SD = 14.20)$ to post-test $(M = 18.84, SD = 15.15), t(42) = 5.04, p <$ $.001\ two - tailed)$.The mean increase in scores for the conc variable was 4.19 with a 95% confidence interval ranging from2.51 to 5.86. Cohen's d statistic (0.77) indicates a medium effect size. The null hypothesis is rejected and we conclude that the teaching intervention contributed to an increase in post scores for the conc variable for experimental group B.

5.3.1.3 Paired Samples Statistics of Control group C pre-test and post-test

Paired Samples Statistics F for control group C

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	CONTROL C SKILL POST	13.514	40	5.3365	.8438
	CONTROL C SKILL PRE	14.242	40	5.8006	.9171
Pair 2	CONTROL C CONC POST	14.00	40	8.181	1.293
	CONTROL C CONC PRE	16.667	40	7.4893	1.1842

Table 5.16 Paired sample statistics of control group C pre-test and post- test

Paired Samples Test for control group C pre to post-test

		Paired Differences					t	df	Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1	CONTROL C SKILL POST - CONTROL C SKILL PRE	-.7279	2.8451	.4499	-1.6378	.1820	-1.618	39	.057	.114
Pair 2	CONTROL C CONC POST - CONTROL C CONC PRE	-2.6667	8.0821	1.2779	-5.2515	-.0819	-2.087	39	.022	.043

Table 5.17 Paired sample statistics of control group C pre-test and post- test

Paired Samples Effect Sizes control group C

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Pair 1	CONTROL C SKILL POST - CONTROL C SKILL PRE	Cohen's d	2.8451	-.256	-.569	.06
		Hedges' correction	2.8728	-.253	-.564	.06
Pair 2	CONTROL C CONC POST - CONTROL C CONC PRE	Cohen's d	8.082	-.330	-.646	-.01
		Hedges' correction	8.161	-.327	-.640	-.00

a. The denominator used in estimating the effect sizes.

Cohen's d uses the sample standard deviation of the mean difference.

Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Table 5.18 Paired Samples Effect Sizes control group C

There was not a statistically significant increase in scores for the variable Control C skill from pre-test $(M = 14.2,\ SD = 5.8)$ to post-test $(M = 13.5, SD = 5.34), t(39) = -1.618, p = .114\ (two - tailed)$. A decrease of -.73 in mean scores was noted for the skill variable with a 95% confidence interval ranging from -1.64 to .018. The Cohen's d statistic (-.569) indicated a negative effect size. The null hypothesis for the skill variable is rejected that is $H_1: \mu_D \neq 0$. We conclude that since there was no teaching intervention for this group the mean scores regressed.

There was a statistically significant decrease in scores for the variable Control C conc from pre- test $(M = 16.67,\ SD = 7.49)$ to post-test $(M = 14, SD = 8.18), t(39) = -2.09, p = .043\ two - tailed)$.The mean decrease in scores for the conc variable was -2.67 with a 95% confidence interval ranging from -5.25 $to - .08$. Cohen's d statistic (-.33) indicates a negative effect size. The null hypothesis is rejected and we conclude that since there was no teaching intervention the scores regressed.

5.3.2 Independent- samples t-test

As indicated previously we also compared mean scores of control and experimental groups post-test (this was done separately for skill and conc scores). This was done to determine if there is a significant difference between control and experimental groups post intervention. Independent samples t-tests was utilized for this purpose. Based on our study design we will have two comparisons: between experimental group A and control group B; and between experimental group B and control group A.

For the independent-samples t-test the null hypothesis is that there is no significant difference after exposure to the teaching strategy. In other words, the mean difference of the control and experimental groups is zero i.e.

$$H_0: \mu_D = 0$$

The alternative hypothesis is that the intervention caused the post-test scores to be higher or lower than the pre-test scores. In other words, the mean difference is not zero:

$$H_1: \mu_D \neq 0$$

The level of significance is set at $\alpha = .05$ for a two-tailed test.

5.3.2.1 SKILL: EXP A VS CONTROL C

The independent-samples t-test statistics for the skill variable of exp A vs. control C is given below. For all the analysis that follows we will check the sig values for Leverne's test. If the value is larger than $.05$ we will use the first row in the independent samples test table otherwise we will use the second row.

Group Statistics Experimental A SKILL vs. Control C SKILL

	GROUP	N	Mean	Std. Deviation	Std. Error Mean
SKILL POST A VS CONTROL C POST	1	43	29.402	19.0397	2.9035
	2	40	14.732	8.1757	1.2927

Table 5.19 Experimental A SKILLPOST Compared to CONTROL C SKILLPOST

Independent Samples Test Experimental A vs. Control C

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
SKILL POST A VS CONTROL C POST	Equal variances assumed	13.440	<.001	4.501	81	<.001	<.001	14.6699	3.2594	8.1847	21.1550
	Equal variances not assumed			4.616	57.852	<.001	<.001	14.6699	3.1783	8.3075	21.0322

Table 5.20 Experimental A SKILLPOST compared to CONTROL C SKILLPOST

Independent Samples Effect Sizes Experimental A SKILLPOST compared to CONTROL C SKILLPOST

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
SKILL POST A VS CONTROL C POST	Cohen's d	14.8375	.989	.529	1.443
	Hedges' correction	14.9767	.980	.524	1.429
	Glass's delta	8.1757	1.794	1.202	2.373

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control group.

Table 5.21 Independent Samples Effect Sizes Experimental A SKILLPOST compared to CONTROL C SKILLPOST

An independent-samples t-test was conducted to compare skill scores between experimental group A and control C. There was a significant difference in scores for exp group A $(M = 29.40, SD = 19.04)$ and control group C $(M = 14.73, SD = 8.18); t(81) = 4.61,$ $p < 001$ $(two - tailed)$. The magnitude of the differences in the means $(mean\ difference = 14.67$ with a 95% confidence interval$[8.31, 21.03]$. The value for the Cohen's d statistic is $.99$ which indicates a large effect size

5.3.2.2 CONC: EXP A VS CONTROL C

The independent-samples t-test statistics for the conc variable of exp A vs. control C is given below.

Group Statistics Experimental A VS Control C

	V1	N	Mean	Std. Deviation	Std. Error Mean
CONC POST A CONTROL C POST	1	42	21.429	14.0870	2.1737
	2	40	13.906	10.4475	1.6519

Table 5.22 EXPERIMENTAL A CONCPOST compared to CONTROL C CONC POST

Independent Samples Test Experimental A VS Control C CONC

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
CONC POST A VS CONTROL C POST	Equal variances assumed	3.257	.075	2.736	80	.004	.008	7.5223	2.7498	2.0501	12.9945
	Equal variances not assumed			2.755	75.544	.004	.007	7.5223	2.7301	2.0843	12.9604

Table 5.23 EXPERIMENTAL A CONCPOST compared to CONTROL C CONC POST

Independent Samples Effect Sizes Experimental A CONCPOST compared to CONTROL C CONCPOST

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
CONC POST A VS CONTROL C POST	Cohen's d	12.4464	.604	.160	1.046
	Hedges' correction	12.5646	.599	.158	1.036
	Glass's delta	10.4475	.720	.255	1.177

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control group.

Table 5.24 Independent Samples Effect Sizes Experimental A CONCPOST compared to CONTROL C CONCPOST

An independent-samples t-test was conducted to compare conc scores between experimental group A and control group C. There was a significant difference in scores for exp group A $(M = 21.43, SD = 14.09)$ and control group C$(M = 13.91, SD = 10.45); t(80) = 2.74, p = .008$ $(two - tailed)$. The magnitude of the differences in the means $(mean\ difference = 7.52$ with a 95% confidence interval$[2.05, 12.99]$. The value for the Cohen's d statistic is $.60$ which indicates a medium effect size

5.3.2.3 SKILL: EXP B VS CONTROL C

The independent-samples t-test statistics for the skill variable of exp B vs. control C is given below. Since in this case the sig value for Leverne's test is less than .05 we will use the second row of the independent samples test table.

Group Statistics Experimental B VS Control C skill

	V1	N	Mean	Std. Deviation	Std. Error Mean
SKILL B POST VS CONTROL C POST	1	43	20.685	10.0385	1.5309
	2	40	11.645	5.7435	.9081

Table 5.25 EXPERIMENTAL B SKILLPOST compared to CONTROL C SKILLPOST

Independent Samples Test Experimental B vs. Control C skill

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
SKILL B POST VS CONTROL C POST	Equal variances assumed	11.239	.001	4.986	81	<.001	<.001	9.0407	1.8133	5.4329	12.6485
	Equal variances not assumed			5.079	67.729	<.001	<.001	9.0407	1.7800	5.4886	12.5928

Table 5.26 EXPERIMENTAL B SKILLPOST compared to CONTROL C SKILLPOST

Independent Samples Effect Sizes Experimental B SKILLPOST compared to CONTROL C SKILLPOST

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
SKILL B POST VS CONTROL C POST	Cohen's d	8.2544	1.095	.630	1.555
	Hedges' correction	8.3318	1.085	.624	1.540
	Glass's delta	5.7435	1.574	1.014	2.122

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control group.

Table 5.27 Independent Samples Effect Sizes Experimental B SKILLPOST compared to CONTROL C SKILLPOST

An independent-samples t-test was conducted to compare skill scores between experimental group B and control group C. There was a significant difference in scores for exp group B $(M = 20.69, SD = 10.04)$ and control group C$(M = 11.65, SD = 5.74); t(81) = 5.08, p < .001$ $(two - tailed)$. The magnitude of the differences in the means $(mean\ difference =$ 9.04 with a 95% confidence interval$[5.49, 12.59]$. The value for the Cohen's d statistic is 1.095 which indicates a large effect size.

5.3.2.4 CONC: EXP B VS CONTROL C

The independent-samples t-test statistics for the conc variable of exp B vs. control A is given below.

Group Statistics EXP B VS CONTROL C CONC

	V1	N	Mean	Std. Deviation	Std. Error Mean
CONC EXP B VS CONTROL C POST	1	43	18.84	15.152	2.311
	2	40	13.50	11.447	1.810

Table 5.28 EXPERIMENTAL B CONCPOST compared to CONTROL C CONCPOST

Independent Samples Test EXP B VS CONTROL C CONC

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
CONC EXP B VS CONTROL C POST	Equal variances assumed	1.279	.261	1.800	81	.038	.076	5.337	2.965	-.561	11.236
	Equal variances not assumed			1.818	77.806	.036	.073	5.337	2.935	-.506	11.181

Table 5.29 EXPERIMENTAL B CONCPOST compared to CONTROL C CONCPOST

Independent Samples Effect Sizes Experimental B CONCPOST compared to CONTROL C CONCPOST

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
CONC EXP B VS CONTROL C POST	Cohen's d	13.495	.395	-.041	.829
	Hedges' correction	13.622	.392	-.040	.821
	Glass's delta	11.447	.466	.021	.906

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control group.

Table 5.30 Independent Samples Effect Sizes Experimental B CONCPOST compared to CONTROL C CONCPOST

An independent-samples t-test was conducted to compare conc scores between experimental group B and control group C. There was not a significant difference in scores for exp group B $(M = 18.84, SD = 15.15)$ and control group C$(M = 13.50, SD = 11.45); t(81) = 1.80, p = .076\ (two - tailed)$. The magnitude of the differences in the means $(mean\ difference = 5.34$ with a 95% confidence interval$[-.56, 11.24]$. The value for the Cohen's d statistic is $.395$ which indicates a small effect size.

5.4 Conclusions of statistical analysis

This research utilized control and experimental groups with a content of both tests were based on Fractions which is prescribed by CAPS. The aim of this investigation was to determine how exposure to an interleaved distributed intervention practice technique affected learner test performance. Three groups were used in the study. Group A and Group B both of these groups were used as an experimental and control groups whereas the third Group C was used only as a control group.

A paired sample T-tests was performed to compare the means of the pre and posts tests of experimental group A, experimental group B and the control group C. The test was broken up into two sections, the skill section which measured procedural fluency and the conceptual section which measured conceptual understanding.

The paired T tests for Experimental group A (see table 5.11) revealed that within the skill paradigm, from the pre to post intervention, this group improved by approximately 4.1%. Likewise, the conceptual paradigm improved by approximately 2.6%. Moreover, the two sided P test revealed that the intervention created a significant improvement within the skill paradigm as it had 0.002% value. Whereas the intervention did not significantly enhance the Conceptual paradigm as it had a 0.179%. Cohan's D revealed that the improvement within the skills paradigm the intervention has a medium effect size. Consequently, we can state that the intervention was successful to improve the procedural fluency of the learners within experimental group. However, their conceptual understanding did not improve.

The paired sample t tests for Experimental group B (see table 5.13) revealed that within the skills paradigm, from pre to post intervention, this groups mean by approximately 6.3%. Likewise, the conceptual paradigm improved by approximately 4.2%. The two sided P tests revealed that the significance of the improvement was valid for both the skills and conceptual paradigm. Cohan's D revealed that the improvement revealed that the intervention had a large effect size. Consequently, we can state that within this experimental group the intervention was successful in improving learners' procedural fluency as well as their conceptual understanding.

The paired sample T tests for Control group C (see table 5.15) revealed that within the skills paradigm from pre to post intervention, group C's mean decreased by approximately -0.7%. Likewise the conceptual paradigm decreased by approximately -2.7%. The two sided P tests revealed that there was not a significant decease for the Skills paradigm. However the decrease within the conceptual paradigm roves to be significant. Cohan's D revealed that the effect size was negative. Consequently, we can state that within the control group the learners' procedural fluency as well as their conceptual understanding decreased.

If all findings concerning SKILLPRE, SKILLPOST CONCPRE and CONCPOST are considered, then it is plausible to conclude that knowledge retention, procedural fluency and conceptual understanding was improved for the groups that were exposed to the distributed interleaved practice. Whereas, the group that had mass practice however did not show the same improvement, in fact this group showed regression of the aforementioned variables.

Chapter 6: Discussion

Our experimental classroom study aimed to investigate whether interleaved distributed practice impacts learners' mathematics retention as well as whether interleaved distributed practice would affect learners' conceptual and procedural knowledge. Consistent with our hypothesis, our results provide clear evidence of the benefit of interleaved distributed practice to improve learners overall mathematics retention as well as their procedural and conceptual knowledge.

The experimental group A was able to improve their procedural fluency to a much greater extent than the control group C whose knowledge retention digressed within the mass practice teaching strategy. However, experimental group A's conceptual understanding has not improved. This means that the intervention improved group A's understanding of the procedures of the learning content but they would not be able to implement conceptual understanding. It would be fitting to acknowledge that Experimental group A started the investigation as the group with the least mathematics attainment of the three groups. Experimental group A contains a high population of learners who have previously repeated grades, as well as learners who struggle with overall literacy. Therefore, improving their conceptual knowledge would prove to be a challenge.

It is conceivable that the learners' prior knowledge is crucial for the effectiveness of the interleaved distributed practice as stated in chapter 3 wherein the Cognitive load theory is discussed. For learners to be able to develop within fractions effectively, a high number of elements need to be processed simultaneously. Therefore, we could argue that the working memory of learners with less prior knowledge may be exceeded.

Subsequently, we could state that learner's accuracy would have increased but not their understanding. It can be argued that it is due to the discriminative-contrast hypothesis that the learners procedural flexibility increased. The discriminative-contrast hypothesis assists learners to differentiate between problem types and allows them to practice finding solutions to questions in different contexts. Therefore, it can be presumed that interleaved distributed practice promotes the different dimensions of knowledge, such as how to apply various procedures within different contexts and also when and why to use these procedures (Nemeth et al., 2021). For this reason, we contend that interleaved distributed practice had a positive impact on experimental group A's learner's procedural fluency.

The experimental group B improved in both procedural fluency as well as conceptual knowledge. The experimental group provides clear evidence that interleaved distributed practice has a positive effect on mathematics retention as well as conceptual and procedural knowledge. This means that within this group, learners are able to perform algorithms effectively as well as understand why and when they are using these strategies. It is fitting to acknowledge that group B is the better performing class within the grade 6 co-hort. The learners have a better understanding of mathematics overall, as well as a better work ethic in the classroom.

Following this, it can be said that because Experimental group B had better prior knowledge, they had a more active schema to retrieve from when the intervention was taking place. They were thus able to have more of an active discourse with their prior knowledge. Therefore, this group of learners were able to make deeper connections when "retrieving," so their memory could be strengthened. As a result learners were able to recall more previously learned knowledge. This action speaks to the Retrieval effect Hypothesis discussed in Chapter 3. Frank and Macnamara (2017) stated that retrieval strategies improve efficiency above and beyond what is gained by practicing the procedure alone. Below we provide examples.

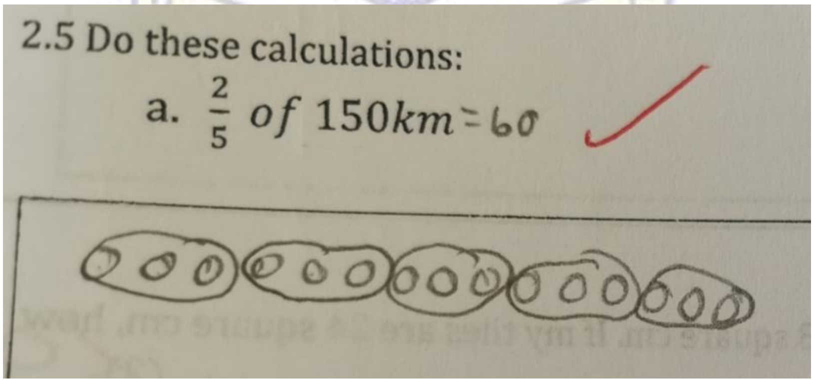


Figure 6.1. Example 1: Finding fractions of whole numbers

In figure 2 the reasoning of a learner trying to find $\frac{2}{5}\ of\ 150$ is shown. The sketch shows that the learner understood that the denominator of 5 was requiring the learner to divide the 150 whole kilometers into 5 equal groups of 10. The tens are represented by three small circles within the bigger ones. Five is then divided into fifteen giving an answer of 3. This 3 is then multiplied by the numerator of 2 giving 6. The learner however kept in mind that each of the little circles represented 10 and the 6 was then multiplied by 10 to give 60. The sketch is an indication of the fact that the learner understood that the operation required was multiplication. It shows that the learner simplified the problem by creating groups of tens. It was further simplified by showing each ten with a single small circle. This allowed the learner to work with smaller numbers making the calculations easier. Creating a visual picture of the required operation also provided a visual aid for reasoning. Creating the figure required the learner to conceptually understand the partitioning of the number 150. In doing the calculation the learner needed to understand the concept that dividing 5 into 150 can be simplified to dividing 5 into 15 since it is an equivalence relation. It is our contention that the teaching strategy allowed the learner to develop a creative approach to the problem by exposing the learner multiple times to the same concepts.

Example 2:

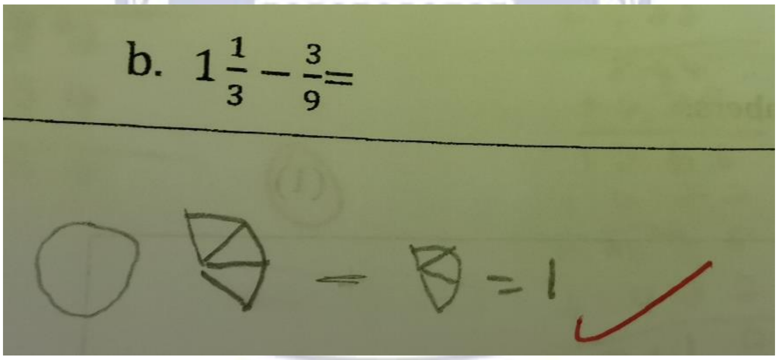


Figure 6.2. Adding fractions with unlike denominators

Figure 3 shows the visual reasoning of a learner doing a mixed fractions subtraction problem. The learner starts by representing the 1 with a full circle. The $\frac{1}{3}$ is represented by 3 parts of a circle. This shows that the learner converted the $\frac{1}{3}$ to $\frac{3}{9}$. Since no calculation is shown the conversion must have been done mentally which is a challenging exercise. This is an indication that the learner understands the concept of equivalent fractions. The to be subtracted $\frac{3}{9}$ is also shown by 3 parts of a circle. The learner therefore solved the problem by changing the problem to an equivalent fractions problem. This shows that the learner have developed a conceptual understanding of equivalent fractions and how to do operations with these. Again it is our argument that the teaching intervention allowed the learner to develop this conceptual understanding through the dual practice types of interleaving and distribution. The distribution allows for more than one encounter with a problem type exposing different facets of the concepts involved. Moreover it requires that students regulary retrieve prior knowledge and engage with it in different ways as determined by the demands of presented problems. This in turn then causes the knowledge to be connected in multiple ways in the memory which provides multiple cognitive hooks to retrieve requisite knowledge. As a consequence retention is improved. Interleaving requires learners to distinguish between problem types and to connect the problem type to the relevant solution procedure. This requires effortful retrieval from long-term memory since the provided information must be compared to stored information and then a selection must be made based on conceptual features of the problem. This process establishes links with prior knowledge but also creates new links. In this way conceptual knowledge is enhanced.

The control group C (under the conventional mass practise system) was however, unable to improve in conceptual understanding or procedural knowledge. In fact, control group C's attainment decreased within both paradigms of our study. Which means that the learners in this group had a better understanding of fractions in the pre-test than in the post-test. The learners were unable to retain more information or even have the same understanding after some time had passed. It is fitting to acknowledge that although our control group seems academically weaker than our experimental groups, they possess quite a good understanding of mathematics overall. Ordinarilly this group would have better attainment than experimental group A.

Therefore, we can state that consistant with our hypothesis our results provide clear evidence of the benefit of interleaved distributed pratice, including positive results for both procedural and conceptual understanding.

Chapter 7: Conclusion

This research was implemented to answer the question on how interleaved distributed practice affects learners' retention of requisite mathematical knowledge. The findings indicate that interleaved distributed practice influenced the study participants 'retention of mathematical knowledge positively. We observed this through the information that the experimental groups (group A and B) were able to retain more relevant knowledge from their pre-test to their post-tests. Whereas the learners who had the mass teaching practice method retain less knowledge from pre-test to post test. This indicates that mass practice helped these learners retain knowledge immediately after a teaching session but that after a delay recall was less efficient. Furthermore, we can also conclude that the interleaved distributed practice not only allowed for retention but conceptual understanding was also enhanced. Whereas, group C who were exposed to mass practice only could not improve on their test scores, in fact their test performance regressed. This means that the mass practice group could only perform well closer to the time where the knowledge was learned but was unable to maintain their knowledge as time progressed. Whereas, interleaved distributed practice allowed learners to improve their retention of mathematics knowledge.

The research also aimed to understand how interleaved distributed practice affected the participants' conceptual and procedural knowledge. In group A, our first experimental group, the learners' procedural knowledge increased from pre-test to post-test. However, their conceptual knowledge did not increase. In group B, our second experimental group, both the learners conceptual and procedural knowledge increased. In group C, our control group, both the learners' procedural and conceptual understanding decreased as a function of time. Therefore, we can state that learners who experience an interleaved distributed practice teaching method have an opportunity to improve their procedural and conceptual understanding. Conversely, learners who are exposed to a mass teaching practice strategy do not have such an opportunity and therefore their conceptual and procedural knowledge decreases.

This study shows that interleaving distributed practice can be successfully implemented in primary school mathematics classrooms. The results of this study can positively influence the common approach to teaching (mass practice) because interleaved can better influence memory retention than massed practice. We have shown in our discussion that participants in our study, did not only rely on standard algorithms, for adding and subtracting fractions and finding fractions of whole numbers, but instead were able to use strategies that were adapted to their own understanding. This indicates that interleaved distributed practice promoted procedural flexibility among learners. Therefore, our results indicate that interleaved practice can promote learners' conceptual understanding as well as their memory of not only fractions but mathematics as a whole. Hence, interleaved practice may also help to promote learners' knowledge in other mathematical topics and domains.

In my experience, CAPS uses a mass teaching practice. This study has shown that learners who utilize a mass teaching practice method perform well only when the knowledge is freshly learnt but as time passes they are unable to perform as well as they did. This clearly shows that mass practice has a clear limitation. Therefore, I recommend that interleaved distributed practice be used as a teaching method.

7.1 Limitations to the study

The study improved procedural knowledge at a more rapid pace than conceptual understanding. Therefore, we could deduce that conceptual understanding takes a longer time to develop as it is the ability to sythesize many different types of knowledge. Hence, the study could have been done over a longer period of time, with more tests inbetween to measure the improvements.

Chapter 8: References

Bjork, R. A. (1994). Memory and metamemory considerations in the. *Metacognition: Knowing about knowing, 185*.

Bruner, J. S. (1966). *Toward a theory of instruction* (Vol. 59). Harvard University Press.

Campbell, D. T., & Riecken, H. W. (1968). Quasi-experimental design. *International encyclopedia of the social sciences, 5*(3), 259-263.

Carey, B. (2015). *How we learn: the surprising truth about when, where, and why it happens*. Random House Trade Paperbacks.

Chen, O., & Kalyuga, S. (2020). Cognitive load theory, spacing effect, and working memory resources depletion: Implications for instructional design. In *Form, function, and style in instructional design: Emerging research and opportunities* (pp. 1-26). IGI Global.

Chen, O., Paas, F., & Sweller, J. (2021). Spacing and interleaving effects require distinct theoretical bases: A systematic review testing the cognitive load and discriminative-contrast hypotheses. *Educational Psychology Review*, *33*(4), 1499-1522

Chiang, I. C. A., Jhangiani, R. S., & Price, P. C. (2015). Quasi-Experimental Research. *Research Methods in Psychology*.

DoE. (2012). Curriculum and Assessment Policy Statement (Intemediate Phase). Retrieved February 1, 2019. from Curriculum and Assessment policy statement (CAPS) Mathematics Gr.4-6 WCED eportal.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4-58.

Feza, N.N. (2014). Keynote address Good intentions are not actions: Mathematics Education of South Africa demands and action and pride from the citizenry.

Foster, N. L., Mueller, M. L., Was, C., Rawson, K. A., & Dunlosky, J. (2019). Why does interleaving improve math learning? The contributions of discriminative contrast and distributed practice. *Memory & cognition*, *47*(6), 1088-1101.

Frank, D. J., & Macnamara, B. N. (2017). Does the acquisition of spatial skill involve a shift from algorithm to memory retrieval?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(12), 1845.

Hu, S. G., Liu, Y., Chen, T. P., Liu, Z., Yu, Q., Deng, L. J., ... & Hosaka, S. (2013). Emulating the Ebbinghaus forgetting curve of the human brain with a NiO-based memristor. *Applied Physics Letters*, *103*(13), 133701.

Hughes, C. A., & Lee, J. Y. (2019). Effective Approaches for Scheduling and Formatting Practice: Distributed, Cumulative, and Interleaved Practice. *TEACHING Exceptional Children*, *51*(6), 411-423.

Kang, S. H. (2016). The benefits of interleaved practice for learning. In *From the Laboratory to the Classroom* (pp. 91-105). Routledge.

May, B. M. (2022). Effects of spaced, repeated retrieval practice and test-potentiated learning on mathematical knowledge and reasoning. *International Journal of Mathematical Education in Science and Technology*, *53*(1), 92-107.

Motshekga, A. (2016). Report of the Ministerial Task Team appointed by Minister Angie Motshekga to investigate allegations into the selling of posts of educators by members of teachers unions and departmental officials in Provincial Education Departments. Department of Basic Education. Final Report 18 May 2016.

Nemeth, L., Werker, K., Arend, J., & Lipowsky, F. (2021). Fostering the acquisition of subtraction strategies with interleaved practice: An intervention study with German third graders. *Learning and Instruction*, *71*, 101354.

Reddy, V., Visser, M., Winnaar, L., Arends, F., Juan, A. L., Prinsloo, C., & Isdale, K. (2016).

TIMSS 2015: Highlights of mathematics and science achievement of grade 9 South African learners.

Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. *Oxford handbook of numerical cognition*, 1118-1134.

Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *Journal of Educational Psychology*, *107*(3), 900.

Ostrow, K., Heffernan, N., Heffernan, C., & Peterson, Z. (2015, June). Blocking vs. interleaving: Examining single-session effects within middle school math homework. In *International conference on artificial intelligence in education* (pp. 338-347). Springer, Cham.

Pan, S. C., Tajran, J., Lovelett, J., Osuna, J., & Rickard, T. C. (2019). Does interleaved practice enhance foreign language learning? The effects of training schedule on Spanish verb conjugation skills. *Journal of Educational Psychology*.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language*, *60*(4), 437-447.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental psychology*, *47*(6), 1525.

Suzuki, Y., Nakata, T., & Dekeyser, R. (2019). The desirable difficulty framework as a theoretical foundation for optimizing and researching second language practice. *The Modern Language Journal*, *103*(3), 713-720.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, *24*(6), 837-848.

Trends In Mathematics Science Studies (n/d). Overview. Retrieved from https://nces.ed.gov/timss/

Western Cape Education Department (2014, October 14). *Background on national quintile system* [Media Release]. Retrieved from: https://wcedonline.westerncape.gov.za/comms/press/2013/74_14oct.html

Western Cape Education Department. (2016, October 08). Research Minute: 0003/2016. Retrieved from https://wcedonline.westerncape.gov.za/circulars/minutes16/EPminutes/edrs3_16.html