

A STUDY OF THE EFFECTS ON RETENTION OF DIFFERENT TIME INTERVALS

BETWEEN OPPORTUNITIES TO LEARN

-----

A thesis submitted in partial fulfilment

of the requirements for the Degree

of

Doctor of Philosophy

in the

University of Canterbury

by

Brett Clark

-----

University of Canterbury

2014

## CONTENTS

LIST OF TABLES	14
LIST OF FIGURES	15
ACKNOWLEDGEMENTS	17
ABSTRACT	19
CHAPTER 1: INTRODUCTION	21
A Short History of Nuthall's Discovery	23
The Nuthall Model and Theories of Learning and Memory	31
Nuthall's model and theories of working memory	31
Nuthall's model and the time window construct	34
Nuthall's model and theoretical accounts of the spacing effect	35
Nuthall's model and neuroscience	38
Previous Research into the Effects on Retention of the Number of Learning Opportunities, Time Windows and Learning Opportunities Distributed in Time	40
Previous cognitive research into the effects on retention of the number of learning opportunities	40
Neurobiological research into the effects on retention of the number of learning opportunities	44
Cognitive research into the effects on retention of time windows	45
Neurobiological research on the time window effect	47
Cognitive research into the effects on retention of the distribution of learning opportunities in time	48
Classroom-based studies into the effects on retention of the distribution of learning opportunities in time	56

Previous neurobiological research into the effects on retention of distributing learning opportunities in time	59
Aims of the Research	61
CHAPTER 2: EXPERIMENTAL DESIGN AND GENERAL METHOD	62
EXPERIMENTAL DESIGN	62
How are Relationships Between Classroom Teaching and Student Learning to be Identified?	68
Control of Extraneous Variables	71
Controlling for the effects of different types of tests	71
Selecting an appropriate retention interval	71
Selecting an appropriate criterion of performance	73
Controlling for difficulty	74
Controlling for content difficulty	74
Controlling for activity difficulty	75
Controlling for the effects of prior learning	77
Controlling the nature of the learning opportunities	78
Measures to control the number of learning opportunities	78
Controlling for the effects of different types of learning outcomes	79
What is the gap between learning opportunities to be called?	80
Controlling the inter-study interval between learning opportunities	81
Controlling the set size	82
Controlling teacher behaviour to enhance treatment fidelity	82
Controlling for carryover effects in the procedural experiments	83
GENERAL METHOD	83

Participants	83
Ethical Considerations	84
Informed Consent of the Participants and Their Parents	84
Teaching Setting	85
EXPERIMENTAL PROCEDURES PART 1	86
Lesson Content of the Experiments Involving the Learning of New Knowledge Items	86
Testing Procedures	86
Pre-tests	86
Retention tests	87
Teaching Procedures and Learning Activities	87
Learning Activity 1	87
Learning Activity 2	88
Learning Activity 3	88
Learning Activity 4	88
Feedback Procedures for the Learning of New Knowledge Items	89
Procedures for Controlling the Number of Learning Opportunities	89
EXPERIMENTAL PROCEDURES PART 2	90
Pilot Study for the Experiments Involving the Learning of New Procedures	90
Pre-teaching for the Operations Experiments	91
Testing Procedures for the Operations Experiments	92
Pre-tests	92
Retention tests	92
Teaching Procedures and Learning Activities	92
Learning Activity 1	94

Learning Activity 2	94
Learning Activity 3	94
Learning Activity 4	94
Feedback Procedures	95
Procedures for Controlling the Number of Learning Opportunities	95
CHAPTER 3: EXPERIMENT 1: EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES FOR UNKNOWN KNOWLEDGE ITEMS IN SETS OF YEAR 7 SCIENCE CLASSES	96
AIM	96
METHOD	96
Participants and Setting	96
Lesson Content	97
Testing Procedures	97
Teaching Procedures and Learning Activities	97
Experimental Treatments	98
RESULTS	99
Inter-marker Agreement	99
Retention of Previously Unknown Knowledge Items	99
DISCUSSION	101
CHAPTER 4: EXPERIMENT 2: EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES FOR UNKNOWN KNOWLEDGE ITEMS IN SETS OF YEAR 13 ART HISTORY LESSONS.	103
AIM	103
METHOD	104
Participants and Setting	104
Lesson Content	104

Testing Procedures	104
Teaching Procedures and Learning Activities	105
Experimental Treatments	106
RESULTS	107
Inter-marker Agreement	107
Retention of Previously Unknown Knowledge Items	107
DISCUSSION	109
CHAPTER 5: EXPERIMENT 3: EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES DURING THE TEACHING OF A NEW MATHEMATICS PROCEDURE TO YEAR 12 STUDENTS	111
AIM	112
METHOD	113
Participants and Setting	113
Lesson Content	114
Pre-teaching	114
Testing Procedures	115
Teaching Procedures and Learning Activities	115
Experimental Treatments	117
RESULTS	123
Inter-marker Agreement	123
Retention of Previously Unknown Procedures	123
DISCUSSION	126
CHAPTER 6: PILOT STUDY: EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES FOR UNKNOWN PROCEDURES IN SETS OF YEAR 13 ART HISTORY LESSONS	128
AIM	128

METHOD	129
Participants and Setting	129
Lesson Content	129
Testing Procedures	129
Teaching Procedures and Learning Activities	130
Pilot Study Treatments	132
RESULTS	137
Inter-marker Agreement	137
Retention of Previously Unknown Procedures	137
DISCUSSION	139
CHAPTER 7: EXPERIMENT 4: EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES BEYOND NUTHALL’S CRITICAL TWO-DAY TIME WINDOW IN SETS OF YEAR 13 ART HISTORY LESSONS	142
AIM	142
METHOD	142
Participants and Setting	142
Lesson Content	143
Testing Procedures	143
Inter-study Intervals	143
Teaching Procedures and Learning Activities	143
Experimental Treatments	144
RESULTS	145
Inter-marker Agreement	145
Retention of Previously Unknown Knowledge Items	146

DISCUSSION	147
CHAPTER 8: EXPERIMENT 5: EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES BEYOND NUTHALL’S CRITICAL TWO-DAY TIME WINDOW IN SETS OF YEAR 12 ENGLISH LESSONS	148
AIM	148
METHOD	148
Participants and Setting	148
Lesson Content	149
Testing Procedures	149
Inter-study Intervals	149
Teaching Procedures and Learning Activities	149
Experimental Treatments	150
RESULTS	151
Inter-marker Agreement	151
Retention of Previously Unknown Knowledge Items	152
DISCUSSION	153
CHAPTER 9: EXPERIMENT 6: EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO THREE TO SEVEN DAYS IN SETS OF YEAR 13 ENGLISH LESSONS	155
AIM	155
METHOD	156
Participants and Setting	156
Lesson Content	156
Testing Procedures	157
Inter-study Intervals	157



Teaching Procedures and Learning Activities	157
Experimental Treatments	158
RESULTS	158
Inter-marker Agreement	158
Retention of Previously Unknown Knowledge Items	159
DISCUSSION	160
CHAPTER 10: EXPERIMENT 7: EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO THREE TO SEVEN DAYS IN SETS OF YEAR 12 ENGLISH LESSONS	162
AIM	162
METHOD	162
Participants and Setting	162
Lesson Content	163
Testing Procedures	163
Inter-study Intervals	164
Teaching Procedures and Learning Activities	164
Experimental Treatments	164
RESULTS	165
Inter-marker Agreement	165
Retention of Previously Unknown Knowledge Items	165
DISCUSSION	166
CHAPTER 11: EXPERIMENT 8: EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO MORE THAN A WEEK IN SETS OF YEAR 13 ART HISTORY LESSONS	169
AIM	169
METHOD	169

Participants and Setting	169
Lesson Content	170
Testing Procedures	170
Inter-study Intervals	170
Teaching Procedures and Learning Activities	170
Experimental Treatments	171
RESULTS	172
Inter-marker Agreement	172
Retention of Previously Unknown Knowledge Items	172
DISCUSSION	174
CHAPTER 12: EXPERIMENT 9: EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO MORE THAN TWO WEEKS IN SETS OF YEAR 13 ENGLISH LESSONS	175
AIM	175
METHOD	176
Participants and Setting	176
Lesson Content	176
Testing Procedures	177
Inter-study Intervals	177
Teaching Procedures and Learning Activities	177
Experimental Treatments	178
RESULTS	179
Inter-marker Agreement	179
Retention of Previously Unknown Knowledge Items	179
DISCUSSION	180

CHAPTER 13: EXPERIMENT 10 EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES BEYOND TWO DAYS DURING THE TEACHING OF A NEW MATHEMATICS PROCEDURE WITH YEAR 10 STUDENTS	182
AIM	182
METHOD	183
Participants and Setting	183
Lesson Content	183
Testing Procedures	183
Inter-study Intervals	184
Teaching Procedures and Learning Activities	184
Experimental Treatments	186
RESULTS	190
Inter-marker Agreement	190
Retention of Previously Unknown Procedures	191
DISCUSSION	193
CHAPTER 14: DISCUSSION	195
Overview of the Present Results	196
The number of learning opportunities required to learn and remember new knowledge items	196
The number of learning opportunities needed to learn a new procedure	197
Pilot Study 1: The learning of new procedures in Year 13 Art History	198
The inter-study interval between learning opportunities	200
The AB Single-Case Experimental Design	204
The Reliability of the Experimental Results	204

The Generality of the Experimental Results	206
Comparing the Results of the Present Study with those from Previous Studies	207
Previous studies which explore the effects on retention of the number of learning opportunities	207
Previous studies on the effects on the retention of varying the time window	208
Previous studies of the spacing effect	208
Theoretical Implications of the Results of the Present Research	209
Theories of working memory	209
The time window construct	210
Cognitive and neurobiological theories of the spacing effect	211
Implications for Teachers	212
Implications for Research on Teaching	216
The Trials and Tribulations of Doing Classroom-Based Research	217
Future Research	217
Conclusion	219
REFERENCES	222
APPENDICES	232
Appendix 1: University of Canterbury Human Ethics Approval Letters	232
Appendix 2: Student Information Sheet and Consent Form	235
Appendix 3: Parent Information Sheet and Consent Form	237
Appendix 4: Test Scores for Each Participant for Each Experimental Treatment	240

This thesis is dedicated to my father, Fred Clark.

## LIST OF TABLES

Table 1.	Number of confirmations and disconfirmations of Nuthall's 3 to 4 learning opportunity rule for the recall of unknown knowledge items for the two experiments which tested this rule.	197
Table 2	Number of confirmations and disconfirmations of Nuthall's 3-4 learning opportunity rule for the recall of unknown procedures for Experiment 3 and Pilot Study 1.	199
Table 3	Number of confirmations and disconfirmations of Nuthall's 2-day rule for the recall of unknown knowledge items for the six experiments which tested this rule.	201
Table 4	Number of confirmations and disconfirmations of Nuthall's 2-day rule for the recall of unknown procedures in Experiment 10.	203

## LIST OF FIGURES

Figure 1.	Diagram of the way curriculum relevant experiences are processed in Nuthall's model of long-term working memory.	33
Figure 2.	Percentage of unknown knowledge items correctly recalled by each of the participants following each of the three treatments in Experiment 1.	100
Figure 3.	Percentage of unknown knowledge items correctly recalled by each of the participants following each of the three treatments in Experiment 2.	108
Figure 4.	Percentage of steps performed correctly and in the correct order by each participant in each of the four treatments (parabola – 4 learning opportunities (4LO), circle graph – 3 learning opportunities (3LO), line graph – 2 learning opportunities (2LO), and hyperbola – 2 learning opportunities (2LO)) in Experiment 3.	123
Figure 5.	Percentage of steps performed correctly and in the correct order by each participant in each of the three treatments (perspective grid, fresco, and tempera) in Pilot Study 1.	138
Figure 6.	Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 2-day/5-day/2-day and 3-day/4-day/3-day schedules of inter-study intervals.	146
Figure 7.	Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 4-day/1-day/0-day, 1-day/3-day/1-day, and 3-day/3-day/1-day schedules of inter-study intervals.	152
Figure 8.	Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 3-day/3-day/3-day and 4-day/6-day/7-day schedules of inter-study intervals.	159
Figure 9.	Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 4-day/3-day/4-day and 6-day/7-day/6-day schedules of inter-study intervals.	166
Figure 10.	Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 5-day/6-day/4-day and 8-day/9-day/5-day schedules of inter-study intervals.	172
Figure 11.	Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 7-day/7-day/7-day and 14-day/28-day/20-day schedules of inter-study intervals.	179
Figure 12.	Percentage of steps correctly performed and in the correct order for each participant following the completion of Treatment 1 (parabola) with 1-day/1-day/1-day ISIs, Treatment 2 (circle) with 4-day/3-day/3-day ISIs	191

and Treatment 3 (line) with 11-day/3-day/6-day or 11-day/7-day/2-day ISIs.



## ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of a number of people. First and foremost I want to acknowledge the support of my principal supervisor, Dr. John Church. Dr. Church's dedication, depth of knowledge, and wisdom have helped me to understand what it takes to be an effective teacher. Dr. Church's great sense of humour made the vicissitudes of this particular academic journey very easy to endure. I wish also to thank Dr. Veronica O'Toole for kindly assisting with the supervision of this research project and for her support and encouragement.

I am most grateful to Linda Collins and Warren Williams for the care they took in check marking the results of the students' tests in each of the experiments in this study.

Critical to the completion of this thesis was the support of the principals of the secondary and primary schools which participated in the research project. I wish also to extend my grateful thanks to the teachers who allowed me to conduct experiments in their classrooms and to the students and their parents for permission to use the results of the student tests for my research.

I also wish to acknowledge my partner, Karen, for her support, understanding and forbearance.

This study was conducted during an extraordinary time in Christchurch. The 7.1 magnitude earthquake which struck on 4 September 2010 ushered in a period of unprecedented seismic activity around Christchurch. The fact that this study was completed at all is testament to the tremendous resilience of all those mentioned above who contributed to the study. In particular I need to express my thanks to the resilience of the students who took part in the study.

Finally I wish to acknowledge the work of the late Professor Graham Nuthall and the work of Dr. Adrienne Alton-Lee. Their classroom based research has provided teachers with

a greater appreciation of not only how students learn but also how to improve their learning.

It is to be hoped that the present study can contribute in some small way to the work of

Professor Nuthall and Dr. Alton-Lee.

## ABSTRACT

In the early 1990s Graham Nuthall and Adrienne Alton-Lee developed a model of the learning and remembering process which has profound implications for teaching and learning at all levels. Using their model they were able to predict what selected primary school students would and would not learn and remember from the teaching of a series of Science and Social Studies units and to do so with an accuracy of between 80 and 85 per cent. The Nuthall model states that for a student to learn and remember a new fact or concept he or she needs three to four learning opportunities with the complete set of information needed to learn the new fact or concept, and a gap of no more than two days between any pair of those learning opportunities. It had always been Graham Nuthall's intention to test the model he had developed with Adrienne Alton-Lee in a series of experiments. Tragically, Professor Graham Nuthall died before this was possible. The ten experiments in this thesis put the Nuthall model to the test.

The first two experiments examined Nuthall & Alton-Lee's claim that to learn and remember new factual information students need three to four exposures to each item of to be learned information. When Year 7 (Grade 6) students were given first four, then three and finally two learning opportunities to learn a set of facts about the planets Mars, Jupiter and Saturn respectively, with a single day between each learning opportunity, the results of the students' retention tests were as the model predicted. A second experiment replicated Experiment 1 using Year 13 (Grade 12) Art History students and sets of unknown knowledge items about three Italian Renaissance paintings. The results of all three treatments in Experiment 2 followed the predictions of the Nuthall model.

Experiment 3 and Pilot Study 1 examined whether or not the first component of the Nuthall model could be generalised to the learning of new operations. Experiment 3 provided a disconfirmation of the model when Year 12 (Grade 11) Mathematics students were able to

learn and remember new operations for graphing a linear equation and a hyperbola after only two learning opportunities. Pilot Study 1 showed that the nature of the new operation had a significant impact on the ability of students to learn and remember new operations.

The second claim made by Nuthall and Alton-Lee was that students would fail to learn and remember new knowledge items when the gaps between pairs of learning opportunities went beyond two days. This claim was found not to be supported across a series of six experiments. The first of these experiments with Year 13 Art History students showed that most students were able to learn and remember more than 80 per cent of a set of new knowledge items when there was a gap of more than two days between pairs of learning opportunities. The gap between successive learning opportunities was gradually extended across the next five experiments.

Finally, the results of Experiment 10 provided further disconfirmations of the Nuthall model as gaps greater than 2 days between pairs of learning opportunities failed to prevent a group of Year 10 students learning new operations for graphing first a circle and then a linear equation.

The findings of the present study strongly suggest that in order to assist students to learn and remember essential curriculum knowledge, teachers need to provide four learning opportunities with inter-study intervals of up to nine days between pairs of learning opportunities. In addition, it would appear that activities which are challenging and force students to engage with the to-be-learned material provide more memorable learning opportunities. Finally, the results of this study provide yet more evidence of the wisdom of designing curricula and textbooks which enable students to revisit essential curriculum knowledge several times during the course of the academic year.

## CHAPTER 1

### INTRODUCTION

We invent so many ways in which to explain why students cannot learn: it is their learning styles; it is their right or left brain strengths or deficits; it is their lack of attention; it is their refusal to take their medication; it is their lack of motivation; it is their parents not being supportive; it is because they do not do their work, and so on. It is not that these explanations are wrong (although some are – there is no support for learning styles, for example) or right (parental expectations and encouragement are powerful factors), but the underlying premise of most of these claims is the belief that we, as educators, cannot change the student. It is this belief that is at the root of deficit thinking. The belief that background factors have the strongest influence on learning would be an argument for putting more resources into poverty and home programs rather than into schooling. We *must* consider ourselves positive change agents for the students who come to us...My point is that teachers' beliefs and commitments are the greatest influence on student achievement *over which we have some control...* (Hattie, 2012, p. 22).

“Know thy impact” (Hattie, 2012, p. 169) is John Hattie’s clarion call to teachers. It’s a clarion call I needed to have heard and heeded much sooner. In 2007 the problem was that I hadn’t. And a greater and more pressing problem was that my school had employed a private company to assess teacher effectiveness and the English department in which I worked was next in line for assessment. Having taken Graham Nuthall’s masters level paper, *Research on Teaching*, some years earlier I knew that, at the very least, I had to have clear evidence of student learning if I was going to pass muster. Knowing that the *inspectors* could arrive in any of my five classes at any time, unannounced, I prepared more thoroughly than ever and created the pretests and post-tests I needed in order to be able to show, hopefully, that student learning had indeed taken place (or that I had a means of demonstrating this in the future).

Two class inspections followed. During both I performed and my students behaved beautifully (bless them). Some of my students were selected to be interviewed about my teaching and its impact on their learning. After their interviews they assured me that they had said nice things about me. A week or so later I was interviewed. In the interview I had to explain how I knew that learning was taking place in my lessons. I presented evidence of pre-

tests and post-tests and showed how carefully I had been monitoring my students' progress. A month or two later we were given our reports. I never read mine. I did not need to. While I had reservations about the teacher evaluation process I had already made an important discovery while I was preparing my lessons for the inspection. The post-tests confirmed it. I discovered that I was a much less effective teacher than I thought I was. It was an uncomfortable discovery to say the least. The one good thing that came out of the experience was that it made me determined to try to find ways of becoming a more effective classroom teacher.

In an ideal world a teacher in my position should have been able to access from his or her laptop relevant high quality and up-to-date research on teacher effectiveness. This was not possible in 2007. Even with the advent of Google Scholar, this is far from being the case today because, after having located some research, the problem then becomes one of trying to determine if the research is actually applicable in the real world.

After completing the teacher assessments, my school certainly did its part to assist its teachers to become more effective teachers. Time was dedicated during professional development days to such things as supplying teachers with lists of the characteristics and practices of good teachers, showing teachers the results of John Hattie's analysis of effective teaching strategies, emphasising the need for pre- and post-testing to demonstrate that learning has happened, introducing the SOLO (Structure of Observed Learning Outcomes) – taxonomy (Hattie, Biggs & Purdie, 1996) into classroom programmes, and providing teachers with excellent professional development on how to improve the quality of teacher feedback to students. Despite all of this I still believed that after 100 years of research in teaching there must be something more out there that could help me become a more effective teacher.

One day during a staff meeting I noticed that one of my colleagues had a copy of Graham Nuthall's book, *The Hidden Lives of Learners*. My colleague, a Physics teacher,

spoke very highly of the book and kindly lent it to me. On page 63 I read for the first time about Nuthall's model of the learning and remembering process. It stated that provided a student has at least three separate encounters with the complete set of information needed to understand a new concept then there is an 80 to 85 per cent chance that the student will learn and remember the concept.

Inspired by what I had read I consciously made an effort to provide at least three opportunities for my classes to engage with the new material they were expected to learn and remember. Doing so took time, however. In my senior classes the pressures of meeting deadlines for NCEA (National Certificate of Educational Achievement) internals and covering the syllabus before students left for exam and study leave in November meant that, over time, it became much more difficult to provide a sufficient number of learning opportunities. While I had anecdotal evidence to suggest that using Nuthall's three learning opportunity rule had helped my students to learn and remember new material, I had no hard evidence to show that this was occurring in my classes. Eventually my desire to see whether Nuthall's model of learning and remembering could predict learning in my students led me to undertake the work described in this thesis.

### **A Short History of Nuthall's Discovery**

In 1991 Graham Nuthall and Adrienne Alton-Lee developed a way of predicting what an individual student will learn on the basis of that student's classroom experiences. They found that the content, frequency and timing of a student's classroom experiences were critical factors in predicting what was learned and remembered. The timing of the experiences played an especially significant role. Nuthall and Alton-Lee discovered that "a gap of more than 2 days between each relevant learning experience resulted in the failure of relevant experiences to connect" (Nuthall, 2000a, p. 96). On the basis of their research Nuthall and Alton-Lee created a model of the learning and remembering process which was

“grounded in continuous and detailed observations and recordings” (Nuthall and Alton-Lee, 1993, p. 836) of students’ classroom experiences in naturally occurring lessons. Using this model they found they were able to predict with an accuracy of between 80-85 per cent what students would and would not remember in post-tests “both immediately after the unit and again 8 to 12 months later” (Nuthall, 2000a, p. 95).

Nuthall and Alton-Lee’s journey of discovery began with Adrienne Alton-Lee’s quest to explain how classroom experiences influenced student learning. She was an experienced teacher who was concerned about how little she knew about how classroom activities affected student learning. Under the supervision of Graham Nuthall she designed her PhD investigations to provide continuous data with which she could track the learning experiences of each of three students in an integrated Social Studies unit. Her three student case studies enabled her to identify a number of key pupil experiences that appeared to be closely related to student learning. These included student opportunity to interact with content and opportunities to attend to concrete demonstrations of new concepts. Using pretest responses and records of pupil opportunity to interact with content, Alton-Lee found she was able to predict over 70 per cent of pupil learning outcomes (Alton-Lee, 1984).

Next Nuthall and Alton-Lee conducted a series of three replications of the original study to see if its results could be repeated. Having created a mountain of data which tracked the learning experiences of eleven 9-12 year old students in four classrooms, Nuthall used a period of study leave to try to make sense of his research results. Try as he might over nine months, he discovered that no matter how he analysed the data the results of the original study could not be replicated. Each new study revealed different relationships between student experiences and student learning outcomes (Nuthall, 2005). Frustrated by his inability to identify any factors that consistently linked student experiences to student learning across



the four studies, he decided to trace the history of one child's learning of one concept. In this way he hoped to gain "an insight into what was going on in that student's mind" (Nuthall, 2005, p. 904). Nuthall was able to analyse the history of a child's learning of a particular concept because of Alton-Lee's development of what she referred to as the *concept file*. "For each concept, proposition, principle, or procedure assessed in the outcome test, a 'concept-file' was created for each student that contained records of all experiences the student had that were related to that concept, proposition, principle, or procedure" (Nuthall, 1999b, p. 148). The creation of each student's concept file was the critical first step in the data preparation procedure used in the Nuthall and Alton-Lee studies. It was Nuthall's detailed analysis of these that would eventually lead to the development of his model of the learning and remembering process.

Nuthall analysed every single exposure that a student (Jon) had experienced with a single concept in a study of recent migration to New York and came to a critical understanding that,

learning involves a progressive change in what we know or can do. What creates or shapes learning is a sequence of events or experiences, each one building on the effects of the previous one. This meant that the same experience might have quite different effects depending on where it occurred in the sequence of an individual student's experiences (Nuthall, 2005, p. 904).

Nuthall's analysis of Jon's experience showed that Jon learned a concept when he experienced the "complete set of information" he needed to fully understand the concept on at least three or four different occasions, provided no gap of more than 48 hours occurred between any two of those experiences. His analysis of Jon's data also revealed "the stream of information relevant to a specific concept experienced by each individual student is dispersed and often fragmented. Even in formally structured activities, information is not encountered in neatly packaged, sequentially organised units" (Nuthall, 1999b, p. 164). Nuthall found that

a “complete set of information” could come from a single experience or it could be made up of several fragments of the complete set of information occurring at different times during a unit. Without at least three or four separate experiences of the complete set of information needed to learn and remember a concept, Jon did not learn and remember that concept.

Having established a procedure for determining when a student had been exposed to the three or four complete sets of information needed to learn and remember a concept, Nuthall then applied the same procedure to the other students in Jon’s study and then in the other two studies. It was then that he discovered he had a problem. He found that there were occasions when students failed to learn and remember a new concept even when it was very clear that they had experienced the complete set of information on three or four occasions. Further analysis of the concept files for these students revealed the critical significance of the timing of the learning experiences. Nuthall found that when there was a gap of more than two days between relevant learning experiences students failed to learn and remember the new concept (Nuthall, 2000a). Thereafter Nuthall developed a more sophisticated set of formal analysis rules to establish a “set of categories for coding the logical content of the curriculum-relevant information the students experienced” (Nuthall & Alton-Lee, 1993, p. 814). The process of identifying when the nature and frequency of a student’s exposure to a fact or concept added up to the complete set of information needed to learn a fact or concept was both complex and time consuming. Each item of knowledge in the post-test was expressed as a proposition, i.e. “a statement which describes a relationship between two or more objects, events, or qualities” (Church, 1999a, p. 19). In the post-tests for each of Nuthall’s six studies each “proposition contained the stem of the item and the correct alternative answer” (Nuthall & Alton-Lee, 1993, p. 814). Thus for the proposition “Winds are caused by the heating and cooling of air” used in Study 4, a self-contained Science unit on weather, students were presented in the post-test with :“Winds are caused by”, i.e. the stem of the proposition, and four alternative

answers: “(a) swirling air and dust, (b) the heating and cooling of air, (c) storms and hurricanes, (d) the change in the seasons, and (e) I don’t know” (p. 814). The coder, whose job it was to trace the learning of this proposition for selected students, had to set up an item file for the proposition. The item file, which was of central importance to Nuthall’s data analysis, consisted of every 15 sec interval of time that was relevant to the proposition. Thus any reference to the proposition or any of its key words such as “‘wind,’ ‘cooling,’ ‘heating,’ [or] ‘air’” (p.814) given in any 15 sec interval of time would be included in the item file for the proposition. Separate item files were created for each outcome test item for every student in each case-study.

After having set up the item file, the job of the coder was then to analyse the nature of the information contained in each 15 sec time interval in the item file, using Nuthall & Alton-Lee’s Curriculum Content Coding Categories. There were three different types of codes, a Content Code, a Validity Code, and a Medium Code and within each code there were numerous categories. The coder’s first task was to establish “the logical relationship between...[ the information contained in the 15 sec interval] and the item proposition” (Nuthall & Alton-Lee, 1993, p. 814). To do this the coder had eight different categories in the Content Code to select from. A 15 sec interval that contained the teacher stating hot and cold air produces winds would be coded as a Content Code 1 because it provided information that contained the complete proposition. Another 15 sec interval may have contained simply the mention of a key word related to the proposition and this would be coded as a Content Code 5, i.e. “Mention of keywords and synonyms” (p. 815).

Having established the content code for the information contained in a 15 sec interval the coder’s next task was to ascertain the Validity Code for that information using the five categories of Validity Code. For example, information that was deemed to be correct and relevant to the item content was given a Validity Code 1. Information that was considered to

be confused or ambiguous was given a Validity Code 2 and eliminated from the item file.

Next the coder had to determine the Medium Code, i.e. how the information was conveyed to the student. Medium Codes included such categories as *Spoken*, *Read*, *Heard* and *Seen*. After establishing this, the coder had one final task to complete before the information contained in the 15 sec interval could be deemed to belong in the student's item file for the proposition.

Detailed observational records were kept for each of the students selected for the study. If the student was observed behaving in a manner which prevented attention to the information then that 15 sec interval was withdrawn from the item file.

With the coding of all the information contained in each of the relevant 15 sec intervals contained in the item file for the proposition completed, the coder then had to apply the set of "construct generation" rules Nuthall & Alton-Lee created for determining if a student would learn and remember the proposition. Typically a student would need many more than just three or four encounters with a concept because many of the encounters provided only a fragment of the complete set of information the student needed to be able to learn and remember the concept. For each of the five different knowledge types – *specific information*, *general concept*, *visual*, *word meaning*, and *skill* – the coders had to apply a set of rules which enabled them to not only determine the extent to which a *content-relevant experience* constituted a complete set of information for learning that concept but also how fragments of information could be combined to create a complete set of information.

To determine if a student had learned and remembered that hot and cold air caused wind the coder would have applied the following set of rules for specific information, i.e. facts:

*For specific information items*, the working memory must contain four instances of Code 1 information (Rule A. 1); or three instances of code 1 information plus one instance of Code 2 or 3 information (Rule A. 2); or two instances of Code 1 information plus two instances of code 2 or 3 information (Rule A. 3); or one instance of Code 1 information plus four or more instances of Code 2 or 3

information (Rule A. 4). The exception to these construct generation rules is that Code 1 information that is incorrect (Validity Code 4) is not included and may also counteract another instance of Code 1 occurring in close proximity (Nuthall & Alton-Lee, 1993, p. 817).

Nuthall discovered that his formalised procedures for analysing the item files for each student enabled him to predict which concepts the students would or would not learn and remember with an average of 80 to 85 per cent accuracy. By 1999 Nuthall and Alton-Lee had the results of six descriptive studies of self-contained Science and Social Studies units confirming the predictive accuracy of their model of classroom learning (Alton-Lee & Nuthall, 1990; Alton-Lee, Nuthall & Patrick, 1993; Nuthall & Alton-Lee, 1991, 1993, 1995).

Nuthall was intrigued by the fact that he could make his predictions based purely on the nature, number and timing of the learning experiences of the children in his studies, without ever having to take account of their academic ability. This led him to question the role that academic ability played in producing differences in individual learning outcomes. After detailed analysis of the subjects' learning experiences during each unit Nuthall came to the conclusion that, "Differences in what students learn, and the way they learn, are a function of differences in the way they participate in classroom activities rather than a function of their underlying or prior abilities" (Nuthall, 1999b, p. 213). Nuthall found that the students who learned and remembered more in his studies did so because they were skilled at generating their own learning opportunities in the classroom (Nuthall, 2005). Zimmerman (1990) refers to these learners as self-regulated learners. He argued that self-regulated learners are distinguished by their skillful use of metacognitive strategies, their high levels of motivation and their behaviour. Self-regulated learners use metacognitive strategies to "plan, set goals, organize, self-monitor, and self-evaluate at various points during the process of acquisition" and that this enables them to be "self-aware, knowledgeable and decisive in their learning" (pp. 4-5). Most importantly they "are aware when they know a fact or possess a skill and

when they do not” (p. 4). Self-regulated learners are *self-starters* who not only work tremendously hard and persevere in difficult situations but also “report high self-efficacy, self-attribution, and intrinsic self-interest” (p. 5). Finally, self-regulated learners “select, structure and create environments that optimize learning” (p. 5).

In the various reports in which Nuthall & Alton-Lee developed their model and described the findings from their descriptive studies they varied the precise number of learning opportunities that they claimed were needed in order to learn and remember new concepts. They also varied their description of the nature of these learning opportunities. In a paper delivered to the Annual Meeting of the American Educational Research Association (1992), Nuthall and Alton-Lee wrote that in order to learn and to remember a specific proposition, students “must experience at least three separate episodes during which they attend to that proposition, or a specific combination of episodes during which they attend to the proposition and further information providing reasons and/or examples of the proposition” (p. 20). In a more comprehensive description of their theory a year later in the *American Educational Research Journal*, Nuthall & Alton-Lee (1993) state the precise number and nature of the learning opportunities in their rules for construct generation. For example, in order to learn and to remember a specific item of information a student’s “working memory must contain four instances of Code 1 information” (p. 817). Code 1 information is defined as “information (spoken, printed, visual, etc) that contains the complete item proposition” (p. 815). In his 1999a article in the *Elementary School Journal* Nuthall states first that students need a “minimum of three to four experiences of the same information” (p. 305) to learn and remember a new concept. Later in the same article he states that for a new knowledge construct to be formed and transferred to long-term memory “at least four representations of category 1 (explicit answer) information” (p. 310) are needed. In his most comprehensive account of both his theory and his descriptive studies,

Nuthall (1999b) specifies that to learn and remember a new concept a student needed to engage “with relevant content on three to four distinct occasions (so that at least three or four different representations of content-relevant information come together in the working memory)” (p. 151). Nuthall produced two further lengthy articles in 2000 in which he developed his theory and described his studies in great detail, and in both he states that students need three to four experiences of relevant information in order to learn and remember a new concept (Nuthall, 2000a; 2000b). In his *Teachers College Record* article Nuthall (2005) makes subtle changes both to the number of learning opportunities and to his description of the nature of the learning opportunities by arguing that, in order to learn and remember a new concept, students must have experienced “the equivalent of three sets of the complete set of information needed to understand ...[the] concept” (p. 905).

### **The Nuthall Model and Theories of Learning and Memory**

*Nuthall’s model and theories of working memory.* “Working memory” has traditionally been conceptualized as the “temporary storage of information in connection with the performance of other cognitive tasks such as reading, problem-solving or learning” (Baddeley, 1983, p. 311). The concept of working memory has evolved over the years. Initially it was conceived of as a limited capacity short-term store (STM) which operated between the sensory register and long-term memory (LTM) (Atkinson & Shrifin, 1968). For Baddeley & Hitch (1974), however, working memory was conceived as a multi-component system, comprising of “an attentional control system – *the central executive* – together with two subsidiary storage systems, the *phonological loop* and the *visuospatial sketchpad*” (Baddeley, 2007, p. 7). As in the original models, the capacity of Baddeley & Hitch’s working memory was strictly limited both in terms of the amount of information that could be processed and the time that new information stayed in working memory.

Nuthall could not accommodate his data with this traditional model of working memory or, it seems, the many other models of working memory which existed at the time (Miyake & Shah, 1999). Nuthall needed a working memory which could not only store “representations of experience for periods of about two days” (Nuthall, 1999b, p. 153) but also “accommodate linkages among thousands of different concepts” (p. 166).

According to the model, a representation of each encounter with concept-relevant information is stored in working memory for up to 2 days and then disappears unless it can attach itself to other representations relevant to the same information that are already in working memory or enter working memory within the 2 day ‘time-window’ ...when at least four representations of ...[the complete set of information needed to learn a concept have] accumulated in working memory, a new knowledge construct ...[is] created from the integration of those representations and transferred to long-term memory ( Nuthall, 1999a, p. 310).

In Nuthall’s model a student’s working memory is an especially dynamic phenomenon capable of sorting and organizing the constant stream of often “dispersed, fragmentary, and overlapping” information that he or she encounters each day in the classroom (Nuthall, 1999b, p. 164).

Students must possess some kind of kind of working memory (or set of spaces in working memory) that sort and organize incoming experience on the basis of the connections that are identified between incoming experiences and prior knowledge structures. Furthermore, once new knowledge structures have been created, they do not remain in long-term memory in an inert state. Each of the components from which they were constructed and each of the related knowledge structures to which they are connected are constantly in use in the interpretation of new experiences and the construction of new knowledge constructs (Nuthall, 1999b, p. 171).

Nuthall’s model of working memory is represented schematically in Figure 1.

The formidable capacity of Nuthall’s model of working memory is only made obvious when you consider the most extreme scenario in which a new knowledge construct can be formed using his model. If one allows for a two-day gap between each pair of four separate encounters with a new concept then Nuthall’s model allows for the creation of a new knowledge construct in working memory over a six-day period before it is finally transferred



to long-term memory. Nuthall suggested that in order to cope with the vast number of different concepts which must be held in his model of working memory at any one time, the processes that take place in working memory which lead to the transference of knowledge to the LTM must largely be both automatic and unconscious (Nuthall, 1999b).

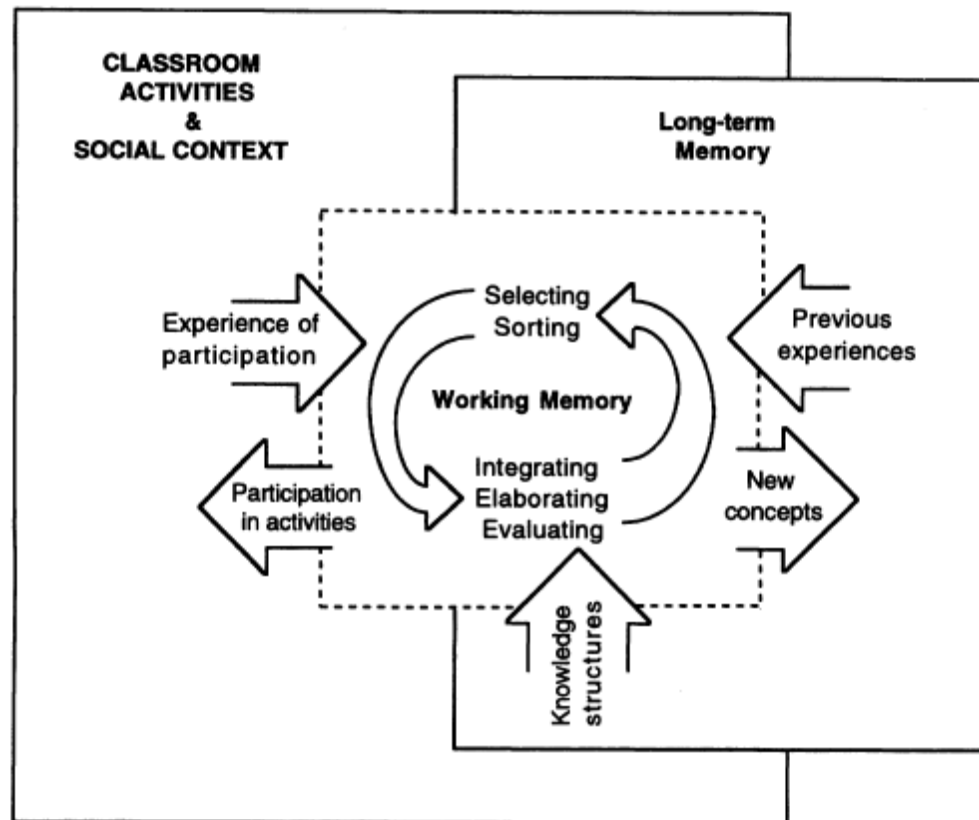


Figure 1. Diagram of the way curriculum relevant classroom experiences are processed in Nuthall's model of long-term working memory (Nuthall, 1999b, p. 151).

Having developed his model Nuthall found support for it in the work of Ericsson & Kintsch (1995). Ericsson & Kintsch (1995) introduced the concept of a long-term working memory (LT-WM), which Nuthall subsequently adopted for his own model, in order to try to account for the ability of expert or skilled performers in specific domains of knowledge to transcend the capacity limits of STM and rapidly access, store and retrieve this domain specific information from LTM. Ericsson & Kinstch's conception of LT-WM bears little

resemblance, however, to that of Nuthall's. The capacity of Nuthall's LT-WM is enormous when compared to the capacity of Ericsson & Kintsch's model of LT-WM.

Baddeley's multi-component model of working memory continued to evolve in response to his own research and the research of others in his field. In order to better account for the way in which working memory and long-term memory interact Baddeley added a fourth component to his model of working memory, the episodic buffer. Baddeley described the episodic buffer as a "temporary storage system that is able to combine information from the [phonological] loop, the [visuospatial] sketchpad, long-term memory, or indeed from perceptual input, into a coherent episode" (2007, p. 148). While the addition of the episodic buffer increased the capacity of Baddeley's model of working memory, its capacity remains very limited when compared to that of Nuthall's long-term working memory. It is interesting to note, however, that while the capacity of Baddeley's episodic buffer and Nuthall's LT-WM are very different, their essential functions appear to be quite similar. Both play an essential role in binding information from many different sources to form what Baddeley terms "chunks", that is "a package of information bound by strong associative links within a chunk, and relatively weak links between chunks" (Baddeley, 2007, p. 148). Nuthall refers to these chunks as "knowledge constructs."

*Nuthall's model and the time window construct.* Nuthall found support for his model of long-term working memory and, in particular, for the critical importance of the two-day time interval between learning opportunities, in Rovee-Collier's research on "time windows" in infant memory. In 1995 Rovee-Collier introduced a new psychological construct into the literature on cognitive development known as the time window. The time window is defined as "the interval between the occurrence of an initial event and the upper limit of the period within which new information can be integrated with the memory representation of that event" (Rovee-Collier, Evancio, & Earley, 1995, p. 69). Rovee-Collier argued "that

information encountered within a given time window will be integrated with information that was encountered previously, but if the same information is encountered outside of the time window, it will not” (Rovee-Collier, 1995, p. 147). For Rovee-Collier, time windows can be seen as “the mortar that holds together the separate building blocks of cognitive development” (p. 166). Nuthall saw Rovee-Collier’s concept of the time window as directly relevant to his work in that Rovee-Collier’s work suggested “that the length of time material is available in working memory is a function of repeated exposures to the same material” (Nuthall, 2000a, p. 97).

*Nuthall’s model and theoretical accounts of the spacing effect.* A great deal of empirical evidence has been produced to demonstrate the efficacy of the spacing effect, that is that learning opportunities distributed over time produce greater recall than learning opportunities which are massed together (Cepeda et al., 2009; Delaney, Verkoeijen & Spiegel, 2010). Many theories have been advanced in an attempt to explain the “spacing effect” but there has been little agreement about the validity of these explanations (Cepeda et al., 2009). There are also difficulties associated with many of these theories because of the problematic nature of the data on which the theories are based (Delaney et al., 2010; Cepeda, Pashler, Rohrer, Vul & Wixted, 2006).

Delaney et al. (2010) argue that their critical analysis of the research on the spacing effect “provides a bleak view of the spacing literature as a whole” (p. 80). They identified five “impostor spacing effects” which can have the effect of either decreasing or increasing any true spacing effect. These are recency effects, rehearsal-borrowing effects on mixed lists, the zero-sum effect on pure lists, deficient processing effects, and list-strength effects. Delaney et al. (2010) argue that “vanishingly few” of the hundreds of spacing studies have controlled for these impostor effects. As a consequence there is a real possibility that many of the theories which try to account for the spacing effect have been developed from data

contaminated by one or more of the impostor effects. “As we... consider the theories of spacing and the evidence against each of those theories, it may be worth keeping in mind that we are using flawed data to reject most of these theories – albeit lots of flawed data collected in multiple laboratories using multiple methods” (Delaney et al., 2010, p. 80).

Cepeda et al. (2009) have also questioned the relevance of many of the theories generated by research into the effects of distributed practice because the vast majority of the research has involved very short inter-study intervals. “Because no one has quantitatively characterized the nature of distributed practice functions over time intervals beyond a day, existing theories of distributed practice may not have much bearing on the phenomenon as it arises over a much longer period” (p. 237). In addition, they argue that the vast majority of the experimental studies of the spacing effect have been limited to verbal list learning in laboratories. Hence existing theories of the spacing effect may not have much bearing on the phenomenon as it manifests itself in real-world classrooms where students are studying authentic material.

Attempts to explain the spacing effect can be grouped under four headings, study phase retrieval theories (Hintzman & Block, 1973), deficient processing theories (Jacoby, 1978), consolidation theories (Landauer, 1969) and Melton’s (1967) encoding variability theory (Vlach & Sandhofer, 2012; Delaney et al., 2010).

*Deficient Processing Theories* attribute the benefits of distributing learning opportunities in time to the reduced level of processing that results when pairs of learning opportunities are close together in time. The idea is that if a student receives a second learning opportunity to learn and remember a new fact or concept shortly after the first learning opportunity, the student’s familiarity with that fact or concept will lead him or her not to invest as much processing time and energy in learning and remembering that fact or concept. Nuthall’s model does not appear to fit comfortably with Deficient Processing

Theories because in Nuthall's model the proximity of the learning opportunities to one another should have no effect on the learning and remembering outcome provided, of course, there is not a gap of more than two days between any pair of learning opportunities. With Nuthall's model it is not a question of the gaps between relevant encounters of new concepts being too small to produce sufficient processing of new knowledge but rather, provided a student experiences the requisite number of relevant encounters with new knowledge, it is a question of whether the size of the gaps between pairs of relevant encounters extends beyond the critical two-day time window.

*Consolidation Theory* states that when a student experiences a new knowledge item a second time a second memory trace for that fact or concept is created which "inherits the state of consolidation of the first occurrence" (Cepeda et al., 2006, p. 369) of that knowledge item. Long inter-study intervals between the first and second learning opportunity should result in higher states of consolidation into long-term memory being inherited by the second memory trace. Provided the inter-study interval is not too long, each subsequent learning opportunity should heighten the state of consolidation of the memory trace still further. Nuthall's model is not wholly compatible with traditional consolidation theories for two reasons. First, according to Nuthall's theory for previously unknown knowledge no memory trace can exist in long-term memory after just one or two experiences of a new knowledge item. Second, the initial consolidation of the memory, that is the creation of the new knowledge construct, takes place not in long-term memory but in long-term working memory and only after this is the construct transferred to long-term memory.

In *Study Phase Retrieval Theory* the second learning opportunity functions as a cue for recalling the memory trace of a knowledge item created by the first learning opportunity. In order to obtain the spacing effect, participants must first be able to recall the memory trace of the knowledge item created by the previous learning opportunity. The longer the inter-

study interval the more difficult it is to retrieve the memory trace. More difficult retrievals are thought to result in greater strengthening of the prior memory trace. As with Consolidation Theory, if the inter-study interval is too long the participant will not be able to recall the prior memory trace of the knowledge item. Study-Phase Retrieval Theory predicts that up to an optimal length, longer inter-study intervals should lengthen the retention interval. Nuthall's model does not sit comfortably with Study Phase Retrieval Theory. In Nuthall's model there is no optimum length of inter-study interval but simply a time window of two days during which a second exposure must occur if memory is to be strengthened. Nuthall's model says nothing about the strength of the memory trace varying according to when the memory was retrieved within the two-day time window.

In simple terms, *Encoding Variability Theory* (Melton, 1967) states that when a new memory trace of a new knowledge item is formed it represents not only the knowledge item but also the context in which the knowledge item was studied. What determines whether a new knowledge item will be recalled is the similarity between the learning opportunity context and the post-test context because the post-test context acts as a cue for retrieval of the newly learned knowledge item. As the inter-study interval between two learning opportunities grows, encoding variability theory predicts that the probability that a newly learned knowledge item will be retrieved grows. This is because it is more likely that the post-test context will be similar to at least one of the learning opportunity contexts. Nuthall's model is compatible with Encoding Variability Theory. Provided there is no more than a gap of two days between any pair of relevant encounters and allowing for the encoding of the new information in long-term working memory prior to being transferred to LTM, Nuthall's model appears to be consistent with Encoding Variability Theory.

*Nuthall's model and neuroscience.* One of the most impressive features of Nuthall's six descriptive studies was the remarkable persistence of his participants' memories.

Although “memory persistence is central in understanding the neurobiology of learning and memory” (Bekinschtein et al., 2007, p. 262), neuroscience, as yet, can tell us little “about the cellular and molecular mechanisms that ...promote the persistence of memory storage” (Medina, Bekinschtein, Cammarota, & Izquierdo, 2008, p. 65). Memory consolidation is a complex phenomenon involving “a complex network of brain systems and parallel molecular events” (Izquierdo et al., 2006, p. 496). Neuroscience has begun the daunting task of trying to determine the “temporal evolution of memory” (Medina et al., 2008, p. 62.) in order to be able to explain why one experience is remembered while another is quickly forgotten. In seeking an answer to this question, neuroscientists have concentrated mainly on the role of the cerebral cortex. However, a few researchers have also begun investigating the role of the hippocampus, despite the “dominant idea that ...[it] appears to have a temporary role in memory storage” (p. 65).

Medina et al. (2008) suggest that the hippocampus has a significant role to play in promoting the persistence of LTM “after LTM is already formed” (p. 66). Bekinschtein et al.’s (2007) research with one-trial inhibitory avoidance training suggests that in the rat hippocampus, protein synthesis and BDNF (Brain Derived Neurotrophic Factor), a protein extensively implicated in memory processing, are required “during a restricted time window around 12 hours after training for persistence of LTM storage, but not for memory formation” (p. 269). In subsequent research, Bekinschtein et al. (2008) found that the BDNF activated memory phase is, by itself, sufficient for forming LTM traces capable of persisting more than one or two days. To ensure the persistence of LTM traces beyond this point Medina et al. (2008) suggest that what is required is the neurobiological equivalent of practice:

We ...propose that LTM storage is achieved by recurrent rounds of consolidation-like protein synthesis-dependent processes and that induction of BDNF expression at this critical time [12 hours after training] is important for the initiation of a cascade of molecular and cellular events that leads to synaptic

remodeling and growth of new synaptic connections which are involved in various forms of LTM (p. 66).

Nuthall's model of learning and memory appears to be consistent with the work of Bekinschtein et al. (2007; 2008).

Neuroscientific research into the effects of sleep on learning and memory also suggests that the hippocampus has a critical role to play in both the formation and persistence of memory. Recent research suggests that declarative memories (memories for facts and events) are reactivated in the hippocampus during slow-wave sleep (SWS) and that "sleep provides an offline mode of processing that leads to gradual incorporation of newly acquired memories into neocortical networks for long-term storage" (Marshall & Born, 2007, pp. 446-447).

Like Bekinschtein et al. (2007), Eckel-Mahan et al. (2008) hypothesised that, in order to provide the requisite levels of protein for a memory to persist, the biochemical pathways underlying memory may require repeated reactivations. Sleep appears to play an important part in inducing repeated reactivations (Rasch & Born, 2007). Eckel-Mahan et al.'s (2008) recent research suggests that the circadian oscillation of mitogen-activated protein kinase (MAPK) "may contribute to the persistence of hippocampus-dependent memories" (p. 1080). Again, Nuthall's model which provides for two night's sleep in the critical two-day time window appears to be consistent with Eckel-Mahan et al.'s hypothesis.

### **Previous Research into the Effects on Retention of the Number of Learning**

#### **Opportunities, Time Windows, and Learning Opportunities Distributed in Time**

*Previous research into the effects on retention of the number of learning opportunities.*

What makes Nuthall's model so attractive for the teacher in real-life classrooms is that it appears to provide a very efficient and simple way of ensuring that students, regardless of their ability, will learn and remember most of the new factual knowledge they are exposed to.



Nuthall's model appears to offer what is tantamount to a Holy Grail of teaching, a standardised procedure for predicting whether or not learning and retention will occur.

Before examining other research into the effects on retention of the number of learning opportunities, it is important that the concept of a learning opportunity is defined precisely. Nuthall was not able to settle on a term for a single contact with a single knowledge item. For example, Nuthall referred to a learning opportunity variously as engaging with "concept-relevant information" (Nuthall, 1999a, p. 151), a "concept-relevant encounter" (p. 152), an encounter "with specific information" (Nuthall, 1999a, p. 306), as an experience "with the same information" (p. 305), and as a separate episode in which a student attends to a proposition (Nuthall & Alton-Lee 1992). In the present research a single learning opportunity is defined as one response to one stimulus followed by feedback. A count of the number of learning opportunities can be used as a measure of the amount of instruction a student has attended to (Heward, 1994).

Nuthall is not the only researcher who has argued that the number of learning opportunities students receive is a critical determinant of whether and how quickly a student will learn and remember new knowledge (Delquadri, Greenwood, Stretton & Hall, 1983). In research involving young children McLay (2004) reports that the 4 year olds in her study of child initiated and adult initiated learning interactions learned and remembered on average about 80 per cent of the names of a set of 20 unknown model animals after just three 3-term learning opportunities with each of the models. Childers & Tomasello (2002) found that 2 ½-year-old children were able to learn and remember a novel noun, a new verb or a new action after just four exposure sessions in a variety of massed and spaced conditions employing different retention intervals. While the children received at least four exposures to the novel nouns, verbs and actions, it cannot be concluded that it took the children at least four exposures to actually learn and remember those new nouns, verbs and actions. In their

experiment it was quite possible for a child to have learnt a new noun, verb or action with fewer exposures.

There is a considerable body of research on novel word learning which suggests that young children have the ability to learn and remember a novel word after a single exposure to the novel word (Spiegel & Halberda, 2011; Waxman & Booth, 2000). This body of research is known as *fast-mapping* research. Fast-mapping refers to the ability of young children to “readily map words to referents in the world and retain these mappings over time, with only minimal exposure” (Vlach & Sandhofer, 2012, Introduction). Vlach & Sandhofer (2012) point out that the nature of the single exposures to new words in fast-mapping research can vary a great deal from study to study: “Fast mapping research often incorporates saliency, repetition, and generation into word learning paradigms...and many paradigms include a reminder trial immediately preceding the retention test” (Do children forget fast-mapping words?, para 3). Given the variable nature of the “exposures” in many novel word learning studies, it is difficult to make meaningful comparisons between this body of research and Nuthall’s descriptive studies.

In research involving elementary school students, McWilliams, in his investigation of variables affecting instructional efficiency, found that his 6-7 year olds required, on average, five learning opportunities to learn and remember a new spelling word (McWilliams, 2006). Axelrod, Kramer, Appleton, Rockett, & Hamlet (1984) found that it took about six learning opportunities for a normally developing 11 year old student to learn and remember a spelling response while Barbetta, Heward & Bradley (1993) found that 8 to 9 year olds classified as developmentally delayed required an average of about 14 trials per word when sight word responses were defined as responding correctly 24 hours after instruction.

Finally, in research involving undergraduate students Cepeda et al. (2009) controlled the number of learning opportunities in a between-groups experiment which examined the

effect of distributed practice on the retention of a set of 23 not-well-known trivia facts. In this experiment, undergraduate students ( $N = 233$ ) were randomly placed in one of six different experimental treatments. The six treatments differed with respect to the size of the gap between the first and the second learning sessions (0-day gap, 1-day gap, 7-day gap, 28-day gap, 84-day gap, and a 168-day gap). Students received three learning opportunities in the first learning session and a further two learning opportunities in the second session. The first learning opportunity of the second session served as a retention test of the students' learning after just one learning session. Students in the 0-day gap and 1-day gap treatments, the only treatments which meet the two-day time window criterion for the Nuthall model, were able to recall, on average, about 97 per cent and 84 per cent respectively of the trivia facts. As with the Childers and Tomasello (2002) experimental study, it cannot be assumed that the students in the Cepeda et al.(2009) experimental study required all three of the learning opportunities in the first session to learn and remember the set of 23 not-well-known facts. It is quite possible that students could have learned some of the facts after just one or two learning opportunities.

In a series of experiments involving undergraduate students, Rawson & Dunlosky (2011) examined the combined effects of initial criterion level and relearning on the retention of the definitions of eight new psychological terms. While Rawson & Dunlosky did not control the number of learning opportunities in any of their learning sessions in their experiments, they did provide information on the average number of trials subjects took to reach the various criterion levels of performance in each of the learning sessions. The fact that the second learning session in their experiments, which occurred two days after the first learning session, served as an interim retention test makes the results of this test of some relevance to the present study. For the first learning session in the first experiment subjects were "randomly assigned to one of four groups, defined by the criterion level for practice

during Session 1 (1, 2, 3, or 4 correct recalls)” (p. 287). The *one correct recall* group took on average 3.5 trials to reach the criterion and averaged about 53 per cent recall in the interim retention test. It took only about four trials on average for the *two correct recall* group to reach their criterion level of performance and they averaged approximately 65 per cent in the interim retention test. For the *three correct recall* group it took just over five trials to reach their criterion level of performance. This group produced the best interim retention test performance averaging about 72 per cent. The final group, the *four correct recall* group, required on average about six trials to reach their criterion level. On the interim retention test, conducted two-days after the initial learning session, they averaged just a little bit less than the three correct recall group. In the second experiment, which was in essence a replication of the first, Rawson & Dunlosky reduced the groups to just two, a *one correct recall* group and a *three correct recall* group. In Experiment 2 the one correct recall group’s performance with respect both to the average number of trials to reach criterion and the average percentage recall on the interim retention test was very similar to that of the *one correct group* in Experiment 1. This was also the case for the *three correct recall* group in Experiment 2.

The small body of research identified above where researchers have controlled the number of learning opportunities provides some limited support for Nuthall’s claim that three to four encounters with “specific information” are required to learn and remember new knowledge items.

*Neurobiological research into the effects on retention of the number of learning opportunities.* Neuroscience still has a great distance to travel before it will be able to describe the neurobiology of human memory. However, new technologies have produced significant advances in the study of animal brains and mammalian brains in particular and this has extended our understanding of the workings of the human brain. One of the most significant breakthroughs in neurobiological research was the discovery of the long term-

potentiation (LTP) effect (Bliss & Lomo, 1973). Long-term potentiation, LTP, is “an enhancement of a chemical synapse between two neurons, lasting from minutes to months or longer. It is typically induced by high frequency stimulation (HFS) or by stimulation protocols with simultaneous pre- and postsynaptic neural stimulation” (Kommeier & Susic-Vasic, 2012, p. 3). The fact that there are a number of similarities between LTP and learning at the behavioural level has led many investigators today to “use LTP as a surrogate in the search for the cellular processes that encode and consolidate memory” (Lynch, Rex, Chen & Gall, 2008, p. 3).

Increasingly, neurological science is concerning itself with the number of “learning opportunities” that are required to create a long-term memory (Kramar, et al., 2012; Gong et al., 2011). Kramar et al., (2012) found that three separate, spaced theta burst trains are required in adult rat hippocampal slices in order to fully saturate LTP within a specific population of synapses. A theta burst train is a burst of magnetic energy which stimulates neuron firing at the synapse. The theta burst train produces “an afferent activity pattern that mimics neuronal firing during learning” (Kramar et al., 2012, Results section, para. 1). If each theta train burst corresponds to a separate learning opportunity then Kramar et al.’s findings fit well with Nuthall’s model of learning and remembering and hint at the possibility that some very fundamental neural phenomena might be involved in creating memories in the mammalian brain.

*Cognitive research into the effects on retention of time windows.* Rovee-Collier was alerted to the significance of the “time window” construct when conducting experiments on infant memory for categories using the mobile conjugate reinforcement task (Rovee-Collier, Greco-Vigorito, & Hayne, 1993). In this task a crib mobile is suspended on a hook above an infant. Suspended from the same hook is a soft ribbon which is attached to the infant’s ankle. The infant learns

to move a crib mobile which is suspended from the same hook by foot-kicking. Later, the ribbon is disconnected, and infants are tested with a stationary mobile that is the same or different from the training mobile. If they kick above operant level, then they are 'saying' that they recognize the test mobile ('same'); if they do not kick above operant level, then they have discriminated it ('different') or forgotten it (Hsu, V. C., 2010, Introduction, para. 3).

Rovee-Collier et al., (1993) trained groups of infants to recognize an alphanumeric category (an A or a 2) displayed on the bottoms and sides of five yellow blocks on a crib mobile that an infant could move by kicking. Only those groups of infants who had been exposed to a moving butterfly wind chime [movement being the common attribute] 0, 3 or 4 days after the conclusion of training showed "significant transfer" (Rovee-Collier, 1995, p. 151). Other groups exposed to the moving butterfly 5 and 6 days after the conclusion of category testing showed no transfer. Their findings led them to postulate that for this experiment there was a time window of 4 days where infants could integrate the new stimulus, the butterfly, with the training event.

In another experiment, four groups of three-month old infants were trained on the mobile conjugate reinforcement task for two sessions. The groups received their second training session 1, 2, 3 or 4 days after the first training session. All four groups were given a retention test eight days after the first training session. The three groups whose second training session occurred within three days of the first showed significant retention while the one group whose second session occurred four days after the first did not. This suggested that there was a time window of three days for this particular task (Rovee-Collier, 1995).

Other research on the time window construct by Rovee-Collier's team of researchers suggests that time windows are task specific (Rovee-Collier, Evancio, & Early, 1995), vary with age (Muzzio, 1994), have "widths" that can be extended not only by increasing the number of memory retrievals (Hayne, 1990) but also by increasing the time intervals between retrievals (Rovee-Collier, Greco-Vigorito & Hayne, 1993), and do not produce uniform

effects, that is “information that is encountered at the end of a time window has a greater impact on the future width of that time window than information that is encountered when the time window first opens” (Rovee-Collier, 1995, p. 148).

Recent systematic research by Hsu (2010) on the effect of the time window construct on retention with infants between the ages of three months and eighteen months confirmed the first tenet of the time window theory, that “there is a limited period after an event occurs...in which a second event can retrieve and be integrated with the memory of the first event” but provided only partial support for the second tenet of the time window construct which states that “when the integration occurs later in the time window, its effects are more enduring” (Hsu, 2010, Abstract ).

The fact that the number of learning opportunities was not controlled in the above research and the very young age of the subjects makes any comparisons with the Nuthall research questionable.

*Neurobiological research on the time window effect.* In their work on the spacing effect and LTP with rat hippocampal brain slices, Kramar et al., (2012) revealed that only widely spaced theta trains produced additional potentiation. They found that a second theta train delivered one hour after the first doubled the level of long-term potentiation whereas a second theta train delivered 10 to 30 minutes after the first produced no additional long-term potentiation. In addition they found that three theta trains each separated by an hour produced a maximal degree of LTP. Kramar et al., (2012) suggested that a time window exists for the additional potentiation in rat hippocampal slices with a second theta train producing additional potentiation if it came in the period 50 to 90 minutes after the first theta train (Lynch et al., 2013).

The age invariant and species invariant nature of the spacing effect strongly suggest that underlying neural mechanisms may best explain this effect (Delaney et al., 2010).

According to Lynch et al. (2013),

the LTP version of the spaced trials effect appears to involve recruitment into the potentiated state of synapses with initially very high plasticity thresholds...the majority of synaptic contacts in adult hippocampus [sic] are not modified by a single train of theta burst but are in some manner primed by the theta train so as to become responsive to a second bout of theta delivered after a long delay (p. 27).

The process by which the first theta train primes high threshold synapses for potentiation takes time as it involves changes to dendritic spine and synapse morphology (Lynch et al., 2013). While the precise nature and timing of these changes are as yet not fully understood, Lynch et al. (2013) propose that the changes involve protein transfer between synapses in the 40-50 minute refractory period. On the surface this refractory period of between 40-50 minutes appears to provide some support for one of the fundamental tenets of the time window construct which states that the retrieval of a “prior memory of ... [an] initial event late in its time window strengthens it (prolongs its retention) more than ... [the retrieval of a ] memory early in its time window” (Hsu, 2010, General Discussion, para 3). For Rovee-Collier (1995), the time window opens immediately after the onset of an event. However, Lynch et al.’s (2013) research suggests that immediately after the onset of an event biochemical changes take place which prevent the integration of a second event with the first.

*Cognitive research into the effects on retention of the distribution of learning opportunities in time.* Research into the effects of the *distribution of practice* or *spacing* on retention is extensive (Cepeda et al., 2006). Most of this research is research into the effects of massed versus spaced practice on retention. This research has produced one of the most robust findings in educational and psychological research and that is that distributed practice produces better retention than massed practice. The *spacing effect* appears to be age invariant.



As we have indicated previously, the spacing effect has been found with infants as young as three months old (Rovee- Collier, 1995); it also appears to persist through to older adulthood (Kornell, Castel, Eich, & Bjork, 2010). There is also growing empirical evidence that the spacing effect is species invariant. Experiments with rats (Kramar et al. 2012), honeybees (Menzel, Manz, Menzel, & Greggers, 2001) and mice (DeZazzo & Tully, 1995) have provided evidence of the spacing effect, adding further weight to the view that fundamental neural processes may account for this effect.

In their meta-analysis of the spacing literature, Cepeda et al. (2006) were highly critical of past quantitative reviews of the effects of distributed practice for having overlooked a major confound present in many well-cited studies. They argue that the great variety of training procedures used in the second and subsequent learning sessions in a number of studies creates confounds which must throw doubt on the conclusions drawn from these studies. In particular they argue that in studies where the second or subsequent learning sessions require participants to train to a criterion of perfect performance for all items “the distribution of practice is confounded with the amount of practice time” (p. 356).

Cepeda et al. (2006) identified just four studies without confounds which investigated the effects on retention of multiple gaps of one day or more with a retention interval of one day or more. The four studies identified by Cepeda et al. were Edwards (1917), Ausubel (1966), Glenberg & Lehmann (1980) and Childers & Tomasello (2002). Unfortunately the Edwards (1917) study could not be accessed.

Ausubel (1966) investigated the effect on retention of a variation in interval between an original learning session and a review session. Ninety-seven students who were either priests or pre-service priests were split into two groups, an early revision group and a delayed revision group. In each of the two learning sessions the groups were required to read a 2500 word passage on Buddhist theology. For the Delayed Revision group the first and second

learning sessions were separated by seven days, while there was only a single day between the first and second learning sessions for the Early Review group. Each group sat a 45-item multiple choice post-test six days after their second learning session. The mean retention scores of both groups were quite similar with the early review group scoring on average 19.14 and the delayed review group scoring 17.87 in the post-test. This led Ausubel to argue that “early and delayed review were not significantly different in enhancing the meaningful learning and retention of school material” (Ausubel, 1966, p. 197). Only the result of the Early Review group is of relevance to the Nuthall model as the gap between the first and second learning sessions falls within Nuthall’s two-day time window. The Nuthall model states that three to four representations of a new knowledge item need to be in long-term working memory before it is transferred to long-term memory and so the Nuthall model would not have predicted that students in the Early Review group would have been able to score as well as they did in their recognition test after a six-day retention interval.

Glenberg & Lehmann (1980) examined the effect of asking 12 university students to recall words from lists of word pairs with an inter-study interval between learning sessions of a week. Each learning session afforded students just one opportunity to learn each of the 58 word pairs. The group which received two learning sessions separated by a week, the spaced-learning group, achieved a higher level of recall after a retention interval of a week (14 %) than did the massed-learning group who experienced two learning sessions concurrently and were then tested one week later (6%). The Nuthall model would not have predicted this result because the gap between the first and second session for the spaced-learning group exceeded Nuthall’s critical two-day time window.

Childers & Tomasello (2002) explored the effects of varying inter-study intervals with thirty-six 2 ½-year-old children who were learning to comprehend and produce new nouns and verbs. In this experiment each child experienced the following four treatments: (a) a

treatment where they were provided with one exposure to a new noun or verb a day for four consecutive days; (b) a second treatment where they were given one exposure per day for four days, but with an inter-study interval of three days between each one; (c) a third treatment where the children received two exposures one day and then three days later another two exposures on that day and (d) a treatment which gave the children two exposures on the first day, four exposures in one session three days after that and then, finally, two exposures a further three days later. For each treatment the children received three retention tests immediately after, 24 hours after, and one week after the final exposure. Childers and Tomasello (2002) found that “for both nouns and verbs, the best learning conditions were the two conditions that took place over 4 days – and for nouns it did not matter how many days intervened between the individual sessions (either 1 day or 3 days)” (p. 973). Children learned just as well to produce new nouns when they experienced one learning session each day for four days as they did when they experienced four learning sessions where each of the learning sessions was separated by three days. What appeared to be the “key variable in these timing-of-exposures conditions ...[was] the number of different days on which children received language models” (p. 973). The Nuthall model is not wholly consistent with the results of this study. It would not have predicted, for example, that the results of the third treatment for the learning of new nouns would have been similar to the results of the first treatment, given that there was a gap of more than two days between the first two exposures and the final two exposures in the third treatment.

In a second study, Childers & Tomasello examined the effect on retention of four novel nouns of varying the gaps between learning opportunities with 20 children aged two years and six months. All children received four exposures to the novel nouns. There were four conditions: 1-day gaps, 2-day gaps, 5-day gaps and 10-day gaps between the four exposures. While varying the gaps had little effect on children’s ability to comprehend new words it had

a significant effect on children's ability to produce a new word. Very few children were able to produce a new noun when the four learning opportunities were separated by 5-day and 10-day gaps which, Childers & Tomasello argue, "provides suggestive evidence that there are gaps wide enough apart to affect learning adversely" (p. 975). Once again the results of this experiment are not wholly consistent with the Nuthall model given that children were able to comprehend a new noun regardless of whether the gap between exposures was 1, 2, 5 or 10 days.

Cepeda et al. (2009) required 215 undergraduate students to learn a set of 40 Swahili-English word pairs. In the first learning session students were required to learn the set of 40 Swahili-English word pairs to a criterion of two correct responses. In a second session, 0-, 1-, 2-, 4-, 7- and 14-days after the first session, the students received just two learning opportunities (as defined in this present research) in a second learning session. Students sat a post-test 10-days after their second learning session. In this experiment only the 0-day gap, the 1-day gap, and the 2-day gap treatments can be accommodated by the two-day time window criterion of the Nuthall model. The average correct responses for 0-day, 1-day and 2-day groups were about 55 per cent, 75 per cent, and 70 per cent respectively.

There are two factors which make a direct comparison between this research and Nuthall's model problematic. First, in this study the number of learning opportunities was controlled in the second learning session but not in the first. Next, the research that gave rise to the Nuthall model was classroom based whereas this research was completed in a laboratory.

Cepeda, Pashler, Rohrer, Vul & Wixted (2008) and Cepeda et al. (2009) carried out a series of experiments exploring the effect on retention of the relationship between the gaps between learning episodes and the retention interval (i.e. the length of time between the last learning episode and the test) in both recall and recognition tests. In the Cepeda et al. (2008)

experiment 1354 subjects of various ages from their Internet Memory Research panel were randomly divided into groups with different gaps and retention interval conditions. Subjects were first divided into four groups with each group having a different retention interval (RI) (7 days, 35 days, 70 days and 350 days). The 7-day (RI) and 70-day (RI) groups were then divided into six different gaps between learning episode groups with the gaps ranging from 0 to 105 days while the 35-day (RI) and 350-day (RI) groups were divided into seven different gaps between learning episode groups again with the gaps ranging from 0 to 105 days. In the first learning session, in which the number of learning opportunities was not controlled, subjects were required to learn 32 “obscure but true trivia facts” (p. 1097) to a criterion of one perfect recall. After their assigned gap each group experienced a second learning session in which each subject was tested twice on each of the 32 facts and given feedback for each response. Then, after their assigned retention interval, each group of subjects sat first a recall test and then a recognition test of each of the 32 facts. In the recall tests the best performing groups were those groups which had a retention interval of just seven days. Regardless of whether the inter-study interval was 1, 2, 7, 21 or 105 days, students recalled over 80 per cent of the 32 facts. The optimum gap between learning sessions for the 7-day retention interval was one day. Students in this group recalled on average about 98 per cent of the 32 facts. What Cepeda et al. (2008) discovered was that each retention interval had an optimal gap for both recall and recognition.

Again, it is difficult to make direct comparisons between these results and the Nuthall results because of the differences in research procedure. However, it is clear that Cepeda et al.’s findings in this experiment are very much at odds with Nuthall’s insistence that the inter-study interval is a time window which cannot exceed two-days.

Cepeda et al. (2009) examined the effect on retention of different gaps between learning episodes with a fixed retention interval (RI) of 10 days. A set of 215 undergraduate

students were divided into six groups with different gaps between learning episodes of 0, 1, 2, 7 and 14 days. Each group of students received two learning sessions separated by their designated gap. In the first learning session subjects were presented with each of the 40 of the Swahili-English word pairs in a random order for a brief period. Thereafter subjects received the first learning session in which the number of learning opportunities was not controlled. Subjects received test-with-feedback trials until they had reached a criterion of two correct responses for each Swahili-English word pair. In the second learning session subjects experienced two test-with-feedback trials for each of the 40 Swahili-English word pairs. The post-test results for the 1-day, 2-day, 4-day, 7-day and 14-day gap groups were very consistent with each of these groups averaging about 70 per cent. However, the optimal gap was one day for the 10 day retention interval as the one-day gap group averaged 75 per cent.

While the fact that the number of learning opportunities was not controlled in the first session, the fact that this research took place in a laboratory setting, and the nature of the to-be-learned material makes comparisons with the Nuthall model difficult, the consistency of the results achieved are, once again, clearly at odds with the Nuthall model.

In a second experiment 233 undergraduate students were required to learn a set of 23 not-well-known facts and a set of 23 names of not-well-known objects. Cepeda et al. (2009) divided the students into six groups with gaps between learning episodes of 0, 1, 7, 28, 84, and 168 days. The first of the two learning episodes for each subject was preceded by a learning trial for both the set of not-well-known facts and the set of not-well-known objects in which each of the two sets of 23 items was presented in the form of a statement. This was followed by the first learning episode which consisted of “three blocks of 23 test-with-feedback trials” (p. 241). In the second learning episode for both the not-well-known facts and the not-well-known objects students completed “two blocks of 23 test-with-feedback trials” (p. 241). The first of the feedback trials in the second learning episode also functioned

as a recall test. Students sat both a recall post-test and a multiple-choice recognition post-test 6 months after the completion of the second learning episode. The highest average recall test score in the six month retention test for the not-well-known facts came from the 28-day gap between learning episodes group. For the recall of not-well-known facts there was very little variation in recall between the 7-day gap, the 84-day gap group and the 168-day gap group. The highest average recall test score in the six month retention test for the not-well-known objects also came from the 28-day gap between learning episodes group.

The fact that the number of learning opportunities was controlled in both learning sessions makes a comparison of this experiment with the Nuthall model more interesting. The Nuthall model states that provided there is not a gap of more than two days between any pair of learning opportunities, it is possible to predict that students will be able to learn and remember between 80 to 85 per cent of the new knowledge items they are required to learn provided they experience at least three or four learning opportunities. In this experiment where the retention interval was six months for all the experimental treatments, students received three learning opportunities in the first learning session and so, according to the Nuthall model, all students should be able to score between 80 to 85 per cent in the first of the feedback trials in the *second learning episode* which functioned as a first test of recall of facts. In their first feedback trial of the second session students in the 7-day gap group did average about 80 per cent. However, the performances of the 28-day group, 84-day group, and 168-day group in the first feedback trial of the second session were well below what the Nuthall model would have predicted. As the gap increased the decline in performance followed the path anticipated by the *curve of forgetting*. An important point to consider here, however, is the nature of the test itself. In the Nuthall studies the post-tests were mainly multi-choice recognition tests and not simply tests of recall. Significantly, in Cepeda et al.'s experiment, the final-test recognition scores of each of the different gap groups were actually

better than what the Nuthall model would have predicted. “Recognition test performance at 0-, 1-, 7-, 28-, 84-, and 168-day gaps was 91..., 95..., 97..., 98..., 95..., and 96... per cent correct..., respectively” (Cepeda et al. 2009, p. 243). The final-test recognition results for the 7-, 28-, 84- and 168-day gap groups are very much at odds with the Nuthall model.

As a result of this series of experiments, Cepeda et al. (2008; 2009) came to several key conclusions about the effect on retention of the relationship between retention interval and inter-study interval. First, any given retention interval has its own optimal inter-study interval between learning opportunities which maximizes retention. Next, any inter-study interval which is either shorter or longer than the optimal inter-study interval will result in suboptimal retention. Finally, “as RI increases, the optimal gap [between learning opportunities] should increase” (Cepeda et al. 2008, p. 1099). Delaney et al. (2010) argue that the work of Cepeda et al. (2009) demonstrates that performance in a retention test of either pure recall or cued recall is “a function of both the retention interval and the lag between repetitions” (p. 67).

*Classroom-based studies into the effects on retention of distributing learning opportunities in time.* The great majority of studies of the spacing effect have been conducted in the laboratory and have required subjects to simply recall word lists or facts. A review of previous research on the effect of distributed practice on the teaching of mathematics by Rohrer & Taylor (2006), for example, produced not a single published, well-controlled experiment which examined how the spacing of practice sessions affected “the retention of a mathematics task that requires more than verbatim recall” (p. 1211).

Concern for the lack of meaningful research on the spacing effect for classroom learning in mathematics led Rohrer & Taylor (2006) to examine the effect on retention of distributing learning sessions for a mathematics procedure in an attempt to determine if the spacing effect holds true for “cognitive tasks that are more conceptually demanding than those that require only verbatim recall” (p. 1210). They set 116 undergraduates from the



University of South Florida the task of learning a procedure for calculating the number of unique orderings (permutations) of given letter sequences where at least one letter was repeated. A pretest revealed that none of the students knew how to perform the procedure. The students were first divided into two groups, the *Massers*, i.e. participants who experienced two learning sessions with the second following immediately after the first, and the *Spacers*, i.e. participants who experienced two learning sessions which were separated by a week. The Massers and Spacers groups were further divided into two groups on the basis of the length of each group's retention interval (RI) to give the following four groups: Massers with a 1 week RI, Massers with a 4 week RI, Spacers with a 1 week RI, and Spacers with a 4 week RI. Each of the four groups experienced a pretest, a training tutorial in which students were taught how to perform a straightforward two-step procedure for determining the unique number of unique orderings in a given letter sequence, two practice sessions in which participants answered five problems and received feedback for each problem, and a final 5 minute testing session in which five problems had to be solved. No feedback was given after the tests. Students in the Spacers group averaged 70 percent correct in the one week test and 64 percent in the 4 week test while the Massers averaged 75 percent in the one week test but could only average 32 percent at the four week test. With five learning opportunities provided for each experimental group in the first session the Nuthall model would have predicted better performances from all four groups in both the one week and four week retention tests. However, it is important to note that Rohrer & Taylor's research involved a very different learning outcome (the learning of a procedure) to those which Nuthall focused on. The results of Rohrer & Taylor's experiment suggests that the Nuthall model may not be generalizable to the learning of procedures.

A second study of the spacing effect in a real-life classroom was conducted by Sobel, Cepeda & Kapler (2011). They taught 39 fifth-grade students from two classrooms the

definitions of English words thought to be new to fifth-grade students. In the massed learning condition the 39 students received two learning sessions with the second following a minute after the completion of the first. In the spaced learning condition the 39 students received two learning sessions a week apart. Each group sat a retention test five weeks after the completion of the second learning session.

The lesson format was the exactly the same for both lessons and for both the massed and spaced conditions. In each learning session, students had to complete a four step process which required students to engage with the definitions in various ways. Students read the definitions with the teacher in step one, wrote down the definitions in step two, listened to the teacher read the definitions and then studied them for a short time in step three, and wrote down the definitions and used the word in a novel sentence in step four.

Children in the massed learning condition averaged just under 10 per cent in the 5-week post-test while students in the spaced condition averaged just over 20 per cent in their 5-week post-test. Children in the spaced learning condition were almost three times more effective in recalling the definitions than the children in the massed learning condition. In securing these results Sobel et al. (2011) demonstrated that the spacing effect which had been found so consistently in laboratory and web-based studies could also be obtained “using educationally relevant spacing intervals between learning episodes and an educationally meaningful retention interval” (p. 765).

It is very difficult to make a comparison of the findings of this study with the Nuthall model because it is difficult to quantify the number of learning opportunities the students received. What the Sobel et al. study does very effectively, however, is to demonstrate what the Nuthall studies have demonstrated clearly, which is the relevance of classroom-based research to the classroom teacher.

*Previous neurobiological research into the effects on retention of distributing learning opportunities in time.* Brain researchers are becoming increasingly interested in discovering the neural machinery involved in producing the spacing effect (Gong et al., 2011; Kornmeier & Susic-Vasic, 2012; Kramar et al., 2012; Lynch, Kramar, Babayan, Rumbaugh, & Gall, 2013; Naqib, Sossin & Farah, 2012). Temporally spaced stimulation trains have been shown to have a similar effect on LTP that temporally spaced learning opportunities have on learning and memory (Kornmeier & Susic-Vasic, 2012). Just as learning opportunities distributed in time produce better memory performance than massed learning opportunities, stimulation trains distributed in time produce longer lasting LTP than massed stimulation trains (Abraham, Logan, Greenwood, & Dragunow, 2002; Zhou, Tao, & Poo, 2003). In addition, just as Cepeda et al. (2009) have shown that there is a non-monotonic relationship between inter-study interval and retention interval, Albensi, Oliver, Taupin, & Otero (2007) have shown that “LTP enhancement is a non-monotonic peak function of the ...spacing interval between stimulations” (Kornmeier & Susic-Vasic, 2012, p. 4). However, the research into how temporally spaced stimulation trains influence LTP is still in its infancy (Kornmeier & Susic-Vasic, 2012; Lynch et al., 2013) and until recently investigators have only produced research showing the effects on LTP of stimulation trains separated by 5 to 15 minutes (Abraham et al., 2002; Gong et al., 2011; Huang & Kandel, 1994; Naqib et al., 2012; Scharf, Woo, Lattal, Nguyen, & Abel, 2002; Zhou et al., 2003). As discussed previously, a recent study by Kramar et al., (2012) employed a long-term potentiation (LTP) stimulation protocol in which the stimulation trains were separated by up to an hour and a half.

In the recent past a number of researchers have voiced their concerns about reading too much into the results of LTP research. Rose (1992) suggested that the LTP phenomenon in rats may be due solely to the lack of stimulation that laboratory rats are exposed to: “Perhaps their hippocampal cognitive mapping capacity is simply desperately starved of use in a lab,

and so responds almost greedily to the novel inputs offered by neurophysiological stimulation of input paths” (p. 239). McGaugh’s (2000) observation that “it remains a hypothesis that the synaptic mechanisms of LTP underlie memory” (p. 250) remains relevant today. In 2007 Abensi et al. pointed out that while “many past studies have attempted to *directly* link LTP with memory encoding...the overwhelming majority of evidence shows *indirect* associations” (2007, p. 3). Albensi et al. echoed Rose’s concern in pointing out the highly artificial nature of LTP research: “... even the most effective [electrical stimulation] protocols trigger a highly synchronous discharge that is unlikely ever to exist in nature” (p. 10).

Research with rat brains (Kramar et al., 2012; Lynch et al., 2013) and chicken brains (Rose, 2000) has suggested the importance of identifying the biochemical changes that take place at the level of the synapse for deepening our understanding of learning and memory in rats and chickens. As yet we lack brain scanning techniques with sufficient resolution to permit us to look at what is happening at the level of the synapse in the human brain.

While the Nuthall’s model is at odds with traditional conceptions of working memory it is consistent with Rovee-Collier’s time window theory, some of the main theories of the spacing effect such as Encoding Variability Theory, and some of the latest neuroscientific hypotheses to do with memory consolidation. Nuthall’s model is quite consistent with previous cognitive and neurobiological research on the number of learning opportunities needed to learn and remember new knowledge items. It is difficult to compare past cognitive research on the time window with Nuthall’s model because much of it involves research with infants where the learning opportunities have not been controlled. Nuthall’s model appears to be much more consistent with recent neurobiological research on the time window. Finally, Nuthall’s insistence on there being a two-day time window between pairs of learning

opportunities appears to be significantly at odds with both classroom and laboratory research on the spacing effect.

Nuthall's research has been published in numerous academic journals and his model of the learning and remembering process was the subject of an entire issue of the *International Journal of Educational Research* in 1999 (Nuthall, 1999b). Nuthall's model continues to be taught to student teachers in Teacher Education Courses in New Zealand. His book, *The Hidden Lives of Learners*, was published by the New Zealand Council for Educational Research in 2007.

### **Aims of Research**

The present research had four aims.

1. The first was to test the predictive validity of Nuthall's model of the learning and remembering process with respect to the number of learning opportunities that students require in order to learn and remember not only unknown knowledge items but also new procedures.
2. The second aim was to test Nuthall's claim that there is a two-day time window which requires that an inter-study interval of no more than two days between any pair of learning opportunities is necessary in order for students to learn and remember new knowledge items and new procedures.
3. If it is found that Nuthall's claim about the critical nature of the two-day time window is false, a third aim was to explore the effects on retention (of both new knowledge items and new procedures) of increasing the inter-study interval between pairs of learning opportunities.
4. A fourth aim was to try to determine if there was in fact a time window for the learning of new knowledge items and new procedures.

## CHAPTER 2

### EXPERIMENTAL DESIGN AND GENERAL METHOD

We have no more reason to say that all psychologists should behave as I have behaved than that they should all behave like R. A. Fisher. The scientist, like any organism, is the product of a unique history. The practices which he finds most appropriate will depend in part upon this history... When we have at last an adequate empirical account of the behavior of Man Thinking, we shall understand all this. Until then, it may be best not to try to fit all scientists into any single mold. (B. F. Skinner, 1959. pp 99-100).

The only way of determining the validity of the Nuthall claims is to test them experimentally. The purpose of the present chapter is to describe the experimental method which was developed for this purpose, that is, to investigate the effect on learning of (a) varying the number of learning opportunities and (b) varying the gap between pairs of learning opportunities.

### EXPERIMENTAL DESIGN

Nuthall and Alton-Lee used a descriptive method which involved making a record for three to four students in each study of everything they saw, heard, said and did during a series of naturally occurring classroom lessons. Nuthall himself acknowledged that the next step was to test his model with a series of classroom experiments (Nuthall, 1999a). As Pressley, Levin & McDaniel (1987) have claimed, “The entire field of curriculum and instruction is at its best when it is an experimental discipline. New procedures are compared to old ones; old techniques are examined to determine if they are all that they are ‘cracked up to be’” (p. 107).

Before discussing the experimental design used in this study it is important to define what an experiment is. Cohen, Manion and Morrison (2011) argue,

The essential feature of experimental research is that investigators deliberately control and manipulate the conditions which determine the events in which they are interested, introduce an intervention and measure the difference it makes. An experiment involves making a change in the value of one variable – called the independent variable – and observing the effect of that change on another variable –called the dependent variable. Using a fixed design, experimental research can be *confirmatory*, seeking to support or not to support a null hypothesis, or

*exploratory*, discovering the effects of certain variables. An independent variable is the input variable, whereas the dependent variable is the outcome variable – the result (p. 312).

The greatest threat to the validity of any experimental research is the uncontrolled extraneous variable. Extraneous variables are factors whose effects in an experiment can “obscure the effects of the independent variable” (Johnston & Pennypacker, 2009, p. 358).

There are two main experimental paradigms in educational research, the between-groups research design, and the single-case experimental design. The great majority of experimental research in education has been carried out using a between-groups design. Gall, Gall & Borg (2010) argue the between groups design “is almost always the preferred experimental design” (p. 303) and Ary, Jacobs & Sorenson (2010) claim that randomised between-group studies “are considered the gold standard for determining ‘what works’ in educational research” ( p. 271). Between-groups designs are regarded by some authors of educational research method textbooks as being capable of producing “true experiments” because subjects are randomly assigned to groups (Ary et al., 2010; Cohen et al., 2011; Creswell 2012).

Single-case experimental research appeals to some researchers in “education and psychology because it is based on an interest in the effectiveness of an intervention for a single, particular individual” (Mertens, 2005, p. 213). Creswell (2012) has argued that, “Without random assignment, this design [the single-case design] is a quasi-experimental rather than an experimental design” (p. 316). Some authors of educational research methods textbooks disagree (e.g., Ary, et al., 2010; Cohen, et al., 2011; Gall et al., 2010; Mertens, 2005). Gall et al. (2010) argue “single-case experiments are rigorous and time-consuming, and often they involve as much data collection as a design involving experimental and control groups. Furthermore, researchers who conduct single-case experiments are just as concerned about issues of internal and external validity as researchers who conduct group experiments”

(p. 310). Sidman (1960) warned against researchers adopting the view that there are “a set of rules that must be followed in the design of experiments” (p. 214). To do so, he argued, would be “disastrous”:

The fact is that *there are no rules of experimental design*...Every experiment is unique. Experiments are made to find out something we do not yet know. If we knew the results beforehand, there would be no point in performing the experiment. In our search for new information we must be prepared at any point to alter our conception of what is desirable in experimental design...Appropriate experimental design cannot be legislated, either by logical or by empirical principles” (pp. 214-215).

Sidman’s argument echoes the view of R. A. Fisher himself. Fisher argued:

A serious consequence of the neglect to make systematically estimates of the efficiency of different methods of experimentation is the danger that satisfactory methods, or methods which with further improvement are capable of becoming satisfactory, may be overlooked, or discarded, in favour of others enjoying a temporary popularity. Fashions in scientific research are subject to rapid changes. Any brilliant achievement, on which attention is temporarily focused, may give a prestige to the method employed, or to some part of it, even in applications to which it has no special appropriateness” (Fisher, 1990, p. 186).

Regardless of the experimental design that a researcher chooses, in the end what matters is whether or not that experimental design is capable of producing valid inferences about the relationship between the variables of a study (Ary et al., 2010). Johnston and Pennypacker (2009) make this point succinctly: “The researcher’s primary objective is to obtain data that truly represent the relationship between the independent variable and the dependent variable. If this is not accomplished, nothing else matters” (p. 247).

The effectiveness of an experiment can be measured by its degree of internal validity and the extent to which its findings have external validity. The concepts of internal validity and external validity refer to “important influences on the legitimacy of experimental inferences” (Johnston & Pennypacker, 2009, p. 336). “Internal validity ...relates to the validity of inferences drawn about the cause and effect relationship between the independent and dependent variables” (Creswell, 2012, p. 303). “An experiment has high internal validity



when changes in the dependent variable are demonstrated to be a function of the independent variable only” (Cooper, Heron & Heward, 1987, p. 234). In essence, internal variability is concerned with the question of experimental control of extraneous variables (Ary et al., 2010).

Campbell and Stanley (1963) identified the following eight main threats to the internal validity of an experimental design: history effects (unforeseen events that may have occurred during the experiment which might have influenced post-test performance); maturation effects (changes that may have occurred within the subjects while the experiment was being conducted); testing effects (taking a test once may influence the subjects’ performance when the test is repeated); instrumentation effects (produced by changes in the measuring instruments during the course of a study); statistical regression effects (the tendency for high and low scorers in a pretest to score closer to the mean in a post-test); selection bias effects (the result of there being differences in the experimental and control groups before the start of the experiment); experimental mortality effects (effects that result from having subjects pull out of the experiment); and, selection-maturation interaction effects (effects that result from one group having a higher rate of maturation than the other).

Cook and Campbell (1979) identified three additional threats to internal validity. These are the experimenter effect (the personal characteristics of the experimenter can influence the results), the Hawthorne effect (when a subject knows they are participating in an experiment it can influence his or her performance), and diffusion effects (effects which result from members of the experimental group communicating details of their experiences to members of the other group or vice versa).

The internal validity of some within-subjects experimental studies is also vulnerable to the threat of carryover effects. Carryover effects “refer to the lingering impact of the effects of one treatment phase on subsequent phases” (Barlow, Nock & Hersen, 2009, p. 91).

Carryover effects constitute a major threat to the internal validity of a single-case experiment as they “can severely limit the inferences one can draw about the controlling effects of the treatment procedures administered” (p. 92).

Achieving a high degree of internal validity is not sufficient, however, for a researcher interested in applied research. Slavin (2002) makes the point that, “Education is an applied field. Research in education should ultimately have something to do with improving outcomes for children” (p. 20). For researchers who are concerned with doing just that the concept of external validity becomes extremely important.

External validity refers to the extent to which the findings of an experimental study “can be generalised to other individuals, settings and time periods” (Gall et al., 2010, pp. 308-309). Between-groups researchers tend to use the term generalisability, while single-case researchers tend to use the term generality when they are discussing the extent to which the results of a study can be applied more generally or more widely than the study itself. A study’s generalisability or generality in applied research is a critical measure of the success of that study (Johnston & Pennypacker, 1993). To have external validity an experimental study must first have a high degree of internal validity:

Tracking down sources of variability is... a primary technique for establishing generality. Generality and variability are basically antithetical concepts. If there are major undiscovered sources of variability in a given set of data, any attempt to achieve subject or principle generality is likely to fail. Every time we discover and achieve control of a factor that contributes to variability, we increase the likelihood that our data will be reproducible with new subjects and in different situations. Experience has taught us that precision of control leads to more extensive generalization of data (Sidman, 1960, p. 152).

Maximising internal validity, however, can reduce the external validity of a study. “A tension always exists between internal and external validity” (Mertens, 2005, p. 126). The more a researcher strives to exercise extremely tight, laboratory-like controls over extraneous variables in his or her experiment “the less realistic and generalizable it becomes” (Gay,

Mills & Airasian, 2009, p. 243). However, the more a researcher tries to achieve a truly realistic context for his or her experiment the less control he or she is able to exercise over extraneous variables. “In the final analysis... the researcher must seek a compromise between a highly controlled and highly natural environment” (p. 243).

Researchers who employ randomized group experimental designs are able to demonstrate external validity to the extent that their samples are truly representative of the population from which they have been drawn. While randomized group designs allow for generalisations from a sample to a population they do not allow for generalisations to be made from a sample to an individual (Barlow et al., 2009, Cooper et al., 1987; Gay et al., 2009; Johnston & Pennypacker 2009). Researchers who use a single-case design attempt to control variability due to individual differences by measuring performance prior to the introduction of the independent variable and after the introduction of the independent variable. While this provides excellent control of individual differences it provides very little generality. In single-case experimental designs the generality of research findings can only be determined by replication (Cooper et al., 1987). “Replication means repetition, and repetition pervades every nook and cranny of scientific method” (Johnston & Pennypacker, 2009, p. 241).

Replication can provide us not only with information about the reliability of a study’s findings but also the generality of those findings (Sidman, 1960). “Reliability involves the stability of the methods used to obtain the data” (Barlow et al., p. 129). A study’s findings have high reliability “when changes in the data are due to changes in the behavior being assessed and not to the method used to obtain the data” (p. 129). By demonstrating the reliability of the effects of the independent variable the researcher informs us about how certain we can be of our new-found knowledge. By demonstrating generality the researcher “actually increases our knowledge” (Johnston & Pennypacker, 2009, p. 244).

Sidman (1960) made a distinction between what he termed direct replication and systematic replication. Direct replication can be defined as "...repetition of a given experiment by the same investigator" (p. 73). Sidman made a distinction between intra-subject direct replication, repeating the same experiment with the same subject, and inter-subject direct replication, repeating the same experiment across different subjects. The former informs us about the reliability of the findings whereas the latter informs us about the generality of the findings across subjects.

Systematic replication in applied research is defined "as any attempt to replicate findings from a direct replication series, varying settings, behavior change agents, behavior disorders, or any combination thereof" (Barlow et al., p. 322). A successful systematic replication "not only demonstrates the reliability of the findings but adds to the external validity of the treatment effect by showing that it can be obtained under varying conditions" (Cooper et al., 1987, p. 242). Systematic replication must be seen as an ongoing process in applied research because it is the job of science "to find exceptions to a given principle" (Barlow et al., p. 335).

### **How are Relationships Between Classroom Teaching and Student Learning to be Identified?**

Researchers have sought to discover relationships between classroom teaching and student learning using a great variety of research methods. These include descriptive studies, correlational studies, between-groups experiments and within-subject experiments.

In order to demonstrate that a causal relationship exists between the number or distribution of learning opportunities and retention an experimental analysis is required. "The experiment is the design of choice for studies that seek to make causal conclusions, and particularly evaluations of education innovations" (Slavin, 2002, p. 18). Nuthall, too, was

very aware of the need for experimental analysis in education: “Only experimental trials in classrooms can determine whether there is more to be discovered in our understanding of the learning process” (1999a, p. 339).

Researchers have sought to discover relationships between classroom teaching and student learning using both between-groups experiments and within-subject experiments. While each method is capable of providing important information, only one of these two methods is specifically designed to answer questions about behaviour change (learning) in individual learners over time and that is the single case experiment.

While randomized group designs remain extremely popular in educational research, they can only answer questions about the average effects of teaching procedures on the learning of samples of students. They cannot be used to measure learning in individuals because “such methods do not provide information about how any given individual either did perform or might perform in the future” (Cooper et al., 1987, p. 145). In the real world of teaching what matters is individual performance. Today teachers know that more than ever they are being made accountable for the performances of the individuals in their classes. In the real world Sally’s parents want to know why she is not coping with the extension work in mathematics. In the real world Fergus’ parents want to know why his performance in reading has not improved. In the real world teachers must measure and report the achievements of each individual student – not the mean achievement of their class of students.

Nuthall’s model of the learning and remembering process was the result of an intense study of the learning of individuals. He set himself

the task of finding out first what was true for the individual student learning specific content within a single curriculum area, and to determine progressively, through a sequence of evolving replications with different students and different curriculum areas, both what was general and what was specific about student classroom learning (Nuthall & Alton-Lee, 1993, p. 801).

Nuthall never attempted to average results across a number of learners. Since his model applies to the learning of individual learners, then an experimental method which preserves the learning of individuals is required.

It was decided therefore to employ a single-case experimental design for the experiments in this thesis so that a conclusion could be drawn regarding the effects of different numbers of learning opportunities and different inter-study intervals on the learning and retention of individual students. The design is based on an AB design which is “the core of all within-subject [single-case] comparisons” (Johnston & Pennypacker, 2009, p. 261). “A” represents the baseline condition, while “B” represents the first intervention phase or the first experimental treatment.

The AB design has two important limitations. First, if there was just a single subject in the experimental treatment and there was a change from the baseline condition, the A condition in the AB design, which appeared to be due to the intervention condition, the B condition in the AB design, we do not know if the “change in responding is a reliable outcome of switching between these two conditions” (p. 262). To demonstrate reliability we need to replicate the experiment either with the same subjects or else with a number of subjects. Second, we cannot be certain that any change observed as a result of the introduction of the independent variable was actually due to the independent variable. There is the possibility that “extraneous variables selectively associated with the treatment condition could supplement, or detract from, the effect of the independent variable” (p. 262). Due to its limitations, Johnston & Pennypacker (2009) argue,

[it] is a weak basis for defensible conclusions about the influence of the independent variable because it only shows a single instance of changes in behavior corresponding to changes in conditions. It is an experiment because it measures a dependent variable under matched control and independent variable conditions, but it does not provide the kind of information necessary to conclude that the change in responding is functionally related to the independent variable (p. 263).

However, if the findings of an experiment which employs an AB design can be replicated directly and systematically on numerous occasions an argument for a functional relationship between the independent and dependent variable can begin to be developed. Blampied (1999) makes this point succinctly: “For single case research designs to move from demonstrating coincidence to causal inference...replication is the critical step” (p. 99).

### **Control of Extraneous Variables**

To achieve a high degree of internal validity for a single-case design it is important to identify all the potential extraneous variables and, where possible, to devise strategies for controlling them. This section identifies potential extraneous variables and describes the strategies which were developed to control each of these.

*Controlling for the effects of different types of tests.* Nuthall’s descriptive studies used a combination of multiple choice and constructed response items for both pretesting and post-testing. These two types of test measure different things; the multiple choice test measures recognition whereas the constructed response test measures the ability to recall and construct a correct response. Moreover multiple choice tests introduce two potential confounds, the guessing confound and confounds due to variations in the plausibility of the several distractors. To eliminate these confounds from the present experiments it was decided to use only constructed response test items during both pretesting and post-testing.

*Selecting an appropriate retention interval.* According to Nuthall’s model, provided a student has three to four learning opportunities with test items and there is no inter-study interval of more than 48 hours between any pair of these learning opportunities we can predict “with considerable accuracy (80% to 85%) which test items students will answer correctly both immediately after the unit and again 8 to 12 months later” (Nuthall, 2000a, p. 95). It needs to be noted that there was considerable variability in the retention intervals in

Nuthall's studies. In Nuthall's six studies, students sat their first post-test two to three weeks after the completion of the Science or Social Studies unit. In Study 2 and Study 4 students were interviewed about their responses on the post-test three weeks after the initial post-test. In Study 6 students were interviewed eight months after the completion of the unit. In all six studies students sat a final post-test 8 to 12 months after the completion of the Science and Social Studies units.

When we consider the experiences of the students in the period between the completion of their units of work and their final post-test, it is clear that all the students received at least one and some students received at least two additional two-term (stimulus-response) learning opportunities before sitting their final post-test 8 to 12 months after the completion of the unit. The post-test two to three weeks after the completion of the unit is likely to have made this a significant learning opportunity. Roediger & Karpicke (2006) argue "that testing not only measures knowledge, but also changes it, often greatly improving retention of the tested knowledge" (p. 181). The interview provides another significant learning opportunity. In Studies 2 and 4 students were interviewed individually about a week after their post-tests and in Study 5 about four months after the first post-test. In each interview an unanswered copy of the post-test was put in front of the student. The interviewer asked the student first to give the correct answer to each question in the post-test. Next, the student was asked to recall how he or she had learned the answer, and, finally, the student was asked to recall anything he or she had done or experienced during the unit which was of relevance to his or her answer (Nuthall, 1995). At no time during this interview did the interviewer indicate that the student's responses were correct or not. The interviews were lengthy with each interview taking three to four sessions to complete and lasting up to four hours. Clearly the comprehensive nature of the interview itself would have provided further learning



opportunities, albeit two-term learning opportunities, as students tried to use what knowledge they had to construct their responses to the interview questions.

Variability in the testing procedures used in Nuthall's studies raises the question of what would qualify as an appropriate retention interval for the proposed experiments. For the present study it was considered that a retention test 7 to 10 days following the final experimental lesson would provide a sufficient test of Nuthall's model for three main reasons. First, a 7 to 10-day retention interval is long enough to make forgetting visible. The *curve of forgetting* suggests that much forgetting occurs within the first week of the retention interval (McDonald, 1959). Ebbinghaus discovered that when he tested himself on his recall of lists of nonsense syllables, his recall after practice decreased 42 per cent after 20 minutes, 66 per cent after 24 hours, and 75 per cent after seven days (Ebbinghaus, 1885/1913). A graph of Ebbinghaus' results creates a smooth curve of forgetting, showing a sharp initial decline in recall and then a much more gradual decline in recall over time. Secondly, a retention interval of 7 to 10 days is similar to the two to three week retention interval employed in each of Nuthall's six studies. Thirdly, a retention interval of 7 to 10 days provided an opportunity to undertake and complete a number of experiments without disrupting the teaching and learning programme of the schools which participated in this study.

*Selecting an appropriate criterion of performance.* Nuthall (1999a) found his model of the learning and remembering process enabled him to successfully predict what students would and would not learn in his descriptive studies with an accuracy of between 80 to 85 per cent (Nuthall, 1999a, 2005). Given that in Nuthall's descriptive studies the retention tests included both multi-choice and constructed response or recall items while in the present study the retention tests were solely tests of recall, it was considered that an appropriate criterion of performance for testing the Nuthall model would be 80 per cent for the learning and remembering of both knowledge items and procedural steps.

*Controlling for “difficulty”.* There are several different kinds of “difficulty” over which the experimenter seeks to exercise good controls. There is the “difficulty” which is a function of the task and there is the “difficulty” which is a function of the prior learning of the student. The “difficulty” inherent in the task in this study is of two kinds, content difficulty and activity difficulty.

*Controlling for content difficulty.* In Nuthall’s studies he was not able to control the content of each of his studies because his were descriptive studies. In his studies it was the class teachers who created and determined the content for each study. Content difficulty seemed to play little part in determining what students learned and remembered in Nuthall’s studies. In making his predictions of what students would and would not learn and remember in his descriptive studies, Nuthall observed that “If a student participated in a sufficient number of relevant learning activities, each separated by no more than two days, he or she learned and remembered the concept or idea regardless of academic ability” (Nuthall, 1999b, p. 241). This led him to assert that “there are no differences in the underlying learning process for more and less able students” (p. 212).

Controlling for the content difficulty in the present study presented some challenges because it was imperative not only that the participants did not know they were participating in an experiment but also that their contact with the experimental content be tightly controlled. These two constraints meant that conducting pilot studies to determine the difficulty of the content for each experimental treatment was not possible. Students could not be exposed to the content before the experiment started as this would have provided students with an additional learning opportunity. The only way to test the content difficulty therefore would have been to use a different set of students. This would have added another confound to this study because in order to maintain the integrity of the within-subjects experimental

design it is essential that that all aspects of the experiment must be experienced by the subjects in the experiment.

To try to exercise control of the lesson content across several treatments of a single experiment the teacher tried to ensure that the content was at the appropriate curriculum level for the participants. It needs to be acknowledged, however, that this is a weak control and that replications of the single-case experiment across different participants and across different subjects are required to determine the extent to which content difficulty was in fact controlled in this study.

*Controlling for activity difficulty.* Nuthall (1999a) stressed that the nature of the activities themselves was important for his model of the learning process. The activities which provided the three to four learning opportunities should not involve the simple repetition of the knowledge to be learned and remembered. Approximately one week after the immediate retention test and one week after the long-term retention test (8-12 months later) Nuthall's subjects were interviewed at length about how they had used their memories to answer questions on the outcome test (Nuthall, 2000a). After analysing these interviews Nuthall realised that the context of each learning experience played a large part in assisting students to recall the content needed to answer the outcome test questions. The students' interview responses made it clear to Nuthall that "when students experience a narrow range of classroom activities they rapidly lose the ability to distinguish one activity from another in memory" (p. 337). To assist recall Nuthall suggested that a variety of activities is needed so that students experience the same information in different ways. Nuthall also stressed that the activities themselves needed to be reasonably challenging. Bjork (1994) has demonstrated that recall is enhanced by activities which present learners with *desirable difficulties*. In the present study the experimental activities were designed to require participants not only to engage with the information to be learned in a variety of different ways but also to produce

an “effort from within” (James, 1890/1981, p. 646) in order to overcome these desirable difficulties.

A potential confound exists in the present study in experiments designed to test Nuthall’s claim about the number of learning opportunities required to learn and remember new knowledge. The first of these experiments involved a class of 10 and 11 year-old learners while the second involved a much older group of 17 to 18 year-old learners. In these experiments each learning opportunity is provided by a different activity. If one activity is harder than another then potentially it means that one learning opportunity might not have been as effective as another and this may influence the results of an experimental treatment. To try to control for this possibility the teacher not only attempted as much as possible to use commonly used and authentic classroom activities which were of a similar level of difficulty but also took great care to make sure that each student actually completed each activity in each experimental treatment. In addition the activities were repeated as much as possible from treatment to treatment within an experiment and from experiment to experiment. Prior to the commencement of an activity the teacher first modelled how to complete the activity with non-experimental content and he also informed the students that they could ask him any questions during the activity if they were unclear about what they had to do. The teacher took an additional step with the 10 to 11 year-old learners. After the teacher had modelled how to complete the activity he then required the participants to complete two to three exercises using non-experimental content so that they could practice what the activity required them to do. The teacher then walked around the room checking carefully that each student had completed the exercises correctly before signalling to the class that they could begin the activity. Ultimately, however, the only means of determining the extent to which activity difficulty was controlled in the present study was through the direct and systematic replication of each single-case experiment.

*Controlling for the effects of prior learning.* When we learn, our previously acquired knowledge and learning to learn skills have a significant influence on what will be learned and remembered. “Decades of research on prior knowledge effects have made clear that existing knowledge provides an important foundation for new knowledge, and that the level of existing knowledge largely determines the degree to which knowledge is learned” (Schapiro, 2004, p. 166). Some investigators have attempted to quantify the extent to which an individual’s learning history can explain his or her performance on measures of retention. Dochy, Segers and Buehl (1999) found that student learning histories could explain as much as 42 per cent of the variance in student performance in tests of retention. Schapiro (2004) argues that the failure to account for the effects of prior knowledge “can endanger the validity of learning outcomes research” (p. 161).

Researchers have used a variety of procedures in their attempts to control for the influence of prior learning. Randomised groups experiments attempt to control for prior learning by randomly allocating subjects to either an experimental group which is exposed to the independent variable or a control group which is not. Such a design, however, cannot eliminate contamination by prior learning effects. This is due to the fact that different subjects experience the experimental and control treatments. Because each individual subject has a unique learning history and brings to the experiment different knowledge, skills, and understandings, the impact of the experimental treatment is likely to be different too. “Each subject typically experiences only one of the conditions being compared. This means that comparisons of the effects of the independent variable and control conditions must involve different subjects, which confounds the effects of treatment conditions with between subject variability” (Johnston & Pennypacker, 1993, p. 183).

A second way of controlling for prior knowledge, and the one which Nuthall used in his descriptive studies, is to measure retention in each individual learner (without averaging

across learners) and to measure retention on only those knowledge items which were unknown to each individual prior to instruction. It is this second procedure which was selected for the experiments undertaken in this thesis.

*Controlling the nature of the learning opportunities.* It will be recalled that Nuthall referred to learning opportunities in a number of different ways (Nuthall & Alton-Lee, 1993; Nuthall, 1999b). He is not alone in this. Church (1999a) used the term “learning interactions” while Greer (1994) described these as two-term and three-term contingencies. Behaviour analysts refer to *response opportunities*.

In this thesis the term *learning opportunity* will be used to avoid ambiguity and it will be defined as a three-term interaction which provides the learner with a practice stimulus, the opportunity to respond, and a consequence following that response. A three-term learning interaction provides the complete set of information that the subject needs in order to acquire a new item of knowledge. Some researchers have argued that the three-term learning interaction is the basic unit of instruction (Greer, 2002; Randolph, 2007).

*Measures to control the number of learning opportunities.* As previously noted, the process of identifying precisely the nature and frequency of a student’s exposure to a fact or a concept using Nuthall’s methodology was both complex and time consuming. Because the present study was to take the form of a series of experiments, procedures had to be devised for controlling exactly the number of learning opportunities experienced by each learner with respect to each element of the content of the learning outcome. This required control over the number of learning opportunities which occurred for each element of the content within each experimental treatment. To achieve this, the learning activities and the content for each learning activity were devised prior to the start of each experiment. To limit the participants’ access to the new knowledge both inside and outside the classroom it was decided (a) to use content which the participants only had access to during the experimental lessons and (b) to

provide activities which limited how and when each participant engaged with each item of new knowledge to be learned and remembered.

*Controlling for the effects of different types of learning outcomes.* The new learning which a teacher wishes his or her students to acquire can take a number of different forms. Gagne (1985), for example, has argued that learning outcomes can be classified into one of five major types of knowledge: intellectual skills (discriminations, concrete concepts, defined concepts, rules and higher-order rules), cognitive strategies, verbal information, motor skills and attitudes. Kameenui & Simmons (1990) distinguish between verbal forms (which include simple facts, verbal chains and discriminations), concepts, rule relationships and cognitive strategies. Church (1999b) distinguishes between behaviours and skills, responding to new stimuli, procedures, knowledge items, and generalisations. While researchers may differ regarding the precise nature and number of learning outcomes which need to be distinguished, they have created a strong research base for demonstrating that different learning outcomes require different teaching procedures (Church, 1999b, Englemann & Carmine, 1991; Gagne, 1985; Gagne & Dick, 1983; Kameenui & Simmons, 1990).

Nuthall was not able to exercise control of the learning outcomes in his descriptive studies because the classroom teacher selected both the learning outcomes and the teaching procedures. Nuthall's analysis of the items in his post-tests for each of the six studies revealed that these items involved five different types of content: specific information (facts), general concepts, visuals (maps and diagrams), word meanings and procedural skills. While he argued that, "The amount and type of information required to form an item-relevant specific knowledge construct depends on the type of test item" (Nuthall & Alton-Lee, 1993, p. 816), a reading of the construct generation rules show that he applied much the same construct generation rules to all five types of content. In time this led Nuthall (2005) to assert that,

what a student knows is a coherent body of beliefs and understandings that is for the most part logically and consistently interconnected. Distinctions among facts, concepts, principles, generalisations and procedures exist more in the theories of researchers than they do in the minds of the students (p. 911).

Despite Nuthall's assertion that differentiating between learning outcomes is not warranted, it was decided to make a clear distinction between two different types of learning outcomes (factual knowledge and procedural knowledge) in the present study. Procedures were included because while Nuthall's studies analysed in detail student learning of facts and concepts he did not apply his model to the learning and remembering of procedural skills. In a procedure, such as a mathematical operation, a number of steps have to be performed in the correct order in order to produce a correct answer or outcome. In order to find out whether Nuthall's model applies to the learning of procedures an additional control has to be introduced. The experimenter must take steps to ensure that the student can perform each of the component steps in a procedure before actually teaching the procedure, that is the sequence of steps itself. This is because, if the student fails to acquire a new procedure after four learning opportunities this may be for either of two different reasons: (a) the student failed to acquire one (or more) of the steps or (b) the student failed to acquire the sequence of steps. To measure (b) requires the experimenter to control (a). To test whether Nuthall's model of the learning process holds for learning procedural knowledge, the student's knowledge of the sequence of actions must be the only thing which is tested.

For these reasons it was decided to limit the learning outcomes under examination to two: the learning and remembering of factual knowledge and the learning and remembering of procedures where the component steps are ones which the student can already perform.

*What is the gap between learning opportunities to be called?* What makes Nuthall's model of the learning process unique is the fact that it specifies not only the number of learning opportunities individuals require in order to learn and remember new knowledge but



it also specifies a critical time period between successive learning opportunities, beyond which individuals will fail to learn and remember new knowledge. Researchers have used a variety of terms to refer to the time interval between successive learning opportunities.

Rovee-Collier (1995) refers to these time intervals as time windows. A time window is defined as “the interval between the occurrence of an initial event and the upper limit of the period within which new information can be integrated with the memory representation of that event” (Rovee-Collier et al., 1995, p. 69).

Rawson & Dunlosky (2011) in their work on the effects on retention of distributed practice refer to the intervals between retrieval trials as *lags*. While Nuthall (1999a) refers generally to the time interval between learning opportunities as a *gap*, when he refers specifically to the critical two-day period in his model of learning and remembering he often uses the term *time window* (Nuthall, 2000a, p. 97). In the context of research into the effects of massed versus spaced practice the time interval between successive learning opportunities has been referred to as the *inter-study interval* (ISI) (Cepeda et al., 2006) or the *inter-session interval* (Rohrer & Taylor, 2006). The inter-study interval is defined as “the interval separating different study episodes of the same materials” (Cepeda et al., 2006, p. 354.). In this thesis, the term inter-study interval (ISI) will be used to refer to the time interval between pairs of learning opportunities.

*Controlling the inter-study interval between learning opportunities.* In the real world of the classroom a number of factors combine to present real challenges for the applied researcher in education. In the present study school timetables, the demands of the curriculum, the pressure of internal and external assessment, and unscheduled interruptions to the school day often limited the options available for manipulating the inter-study interval between learning opportunities. This meant that it was not possible to implement in any systematic way schedules of inter-study intervals which were either fixed-interval (e.g. a 3-

day, 3-day, 3-day schedule), expanding (e.g. a 3-day, 6-day, 9-day schedule) or decreasing (e.g. a 5-day, 3-day, 1-day schedule) schedules of inter-study intervals. As a consequence, in the present study, most of the experimental treatments in the experiments which tested Nuthall's 2-day rule had schedules of inter-study intervals which were governed by the school's timetable for a particular content (such as Art History). Even within this constraint, however, it proved possible to design a number of experimental tests of Nuthall's 2-day time window claim. Given that there is research to demonstrate that fixed-interval and expanding intervals of inter-study intervals can influence recall in some contexts in different ways (Roediger & Karpicke, 2011), there is the possibility that in the present study the use of a mixture of fixed-interval, expanding, expanding/decreasing and decreasing/expanding schedules of inter-study intervals may have confounded the results of the seven experiments which tested Nuthall's 2-day rule. The systematic replication of the gap experiments may reveal the extent to which this potential confound influenced the results of any of the "inter-study interval" experiments.

*Controlling the set size.* If an experimental test of Nuthall's model of the learning process is to be undertaken in the classroom, then the content of experimental lessons must be similar to that which is normally contained within a 25 to 50 minute lesson. For the present investigations a 25 minute lesson was selected. Given that each activity with experimental content was to take no more than 25 minutes, it was considered that for the acquisition of new knowledge items a set size of between 12 to 16 knowledge items (plus some non-experimental items) would be appropriate. As far as procedural skills were concerned, it was decided to use procedures which involved between 12 and 16 steps.

*Controlling teacher behaviour to enhance treatment fidelity.* Treatment fidelity (also referred to as treatment integrity) can be "defined as the extent to which the experimental intervention is implemented according to the specifications of the researchers or program

developers” (Gall et al., 2010, p. 308). A potential source of variability in any series of classroom experiments and thus a threat to treatment fidelity is the behaviour of the teachers who are required to implement the experimental procedures. To try to control this possible source of variability the same teacher was used to conduct the experimental lessons in each of the ten experiments in this study. This helped to ensure that there was high degree of consistency in the way in which the experimental lessons were conducted across the ten experiments.

*Controlling for carryover effects in the procedural experiments.* One of the greatest threats to the internal validity of a single-case experimental study is the sequence effect or carryover effect. A sequence effect refers to the “the presence of uncontrolled effects on the target behavior caused by the sequence of conditions employed in the study” (Cooper et al., 1987). Pre- test sensitization, that is the effect of the pretest sensitizing some of the students to certain items of content in the experimental lessons, poses a potential threat to the internal validity of the results in the present study. The fact that much of the content of each experimental lesson in the present study will be tested by the retention test in each experimental treatment is likely to lessen the impact of any possible sequence effects. In addition, the fact that the same pre-testing procedure will be used for every student in every treatment in the present study is likely to mean that any sequence effects are going to be similar for every experimental treatment.

## GENERAL METHOD

### **Participants**

A Decile 10 Christchurch secondary school and a Decile 10 Christchurch primary school provided the participants for the experiments reported herein. In New Zealand, schools are given a decile ranking of 1 to 10 based on the median income levels of residents in the

school's immediate catchment area. This is based on data from the National Census. Decile 1 schools are located in the poorest areas and Decile 10 schools are located in the wealthiest areas of the country. Experiments involved students in nine separate secondary school classes and a single primary school class. All ten classes were made up of participants of mixed ability and mixed gender. The ages of the primary school students ranged from 11 to 12 years and the ages of the secondary school students ranged from 14 to 18 years of age.

### **Ethical Considerations**

In order to maintain the internal validity of each experiment, it was critical that the participants did not know they were participating in an experiment. The University of Canterbury Ethics Committee granted permission for the experiments provided (a) informed consent was granted by each participating school prior to the start of the experiments and (b) the informed consent of both the participants and their parents was obtained after each individual experiment had been completed. The approval letters are reproduced in Appendix 1.

### **Informed Consent of the Participants and Their Parents**

On the completion of an experiment the classroom teacher informed the participants that they had participated in an experiment which examined the way they learned and remembered new lesson content. The teacher provided each participant with an information sheet outlining briefly the purpose of the research, the nature of the experiment, and the measures that were being taken both to maintain their anonymity and the confidentiality of their results. Participants were then issued with a consent form on which they could indicate whether or not they wished to allow their results to be included in the results presented in this thesis. The consent form restated the measures that would be taken to ensure the confidentiality and anonymity of their results and also pointed out their right to withdraw

their results from the study at any time without penalty. The information sheet and the consent form are reproduced in Appendix 2.

The teacher also issued each participant with a Parent Information Sheet and a Parent Consent Form to take home to their parents. The Parent Information Sheet and the Parent Consent Form are reproduced in Appendix 3. The Parent's Information Sheet outlined the purpose of the research, the nature of the experiments and the measures that were being taken to safeguard their child's confidentiality and anonymity. It also pointed out that during the experiment their son or daughter had been learning his/her normal curriculum content using normal teaching practices. The accompanying Parental Consent Form provided a summary of the conditions under which the experiment had been undertaken, indicated again the measures that would be taken to safeguard the confidentiality and anonymity of their child, and pointed out their right to withdraw their child's results at any time without penalty. During the course of the experiments no parent requested their child's results to be withdrawn and no student requested their test results to be withdrawn from the project.

In order to safeguard the anonymity participants were originally to have been referred to by a code number in this thesis but during the preparation of this thesis it was decided to change from code numbers to code names in the interests of clarity. This decision was supported by both supervisors of this study. A lengthy list of commonly used male and female Christian names was generated to create a set of code names. Then each of the participants in each of the ten experiments and the pilot study was allocated a code name at random.

### **Teaching Setting**

Teaching and testing were conducted for the most part in the participants' normal classrooms with their normal seating arrangements. Earthquake damage caused by the devastating Christchurch earthquake on 22 February 2010 meant that many classes found

themselves being timetabled into a variety of different classrooms for extended periods of time. Despite the many changes of classroom all the classrooms used during the course of this study were well equipped and provided good environments for learning. In all but the experiment involving primary school participants the teacher conducting the experiment was the participants' usual classroom teacher.

## EXPERIMENTAL PROCEDURES PART 1

### **Lesson Content of the Experiments Involving the Learning of New Knowledge Items**

Care was taken to select the content for each experiment from the school curriculum. The main reason for this was to ensure high levels of external validity. For each 50 minute experimental lesson in each experimental treatment participants were required to complete two tasks. Task 1 always contained the experimental content and Task 2 always contained non-experimental content. Given that each activity with the experimental content was expected to take between 20 and 25 minutes a set size of between 12 and 16 knowledge items was considered to be within the capabilities of the participants and this was the set size for each experiment.

### **Testing Procedures**

*Pre-tests.* As in the Nuthall studies, knowledge items were expressed as two-element propositions and test items were constructed by turning one element of the proposition into a question. For example the test question for the proposition "Saturn was the Roman God of agriculture and harvests" took the form "What was Saturn the Roman God of?"

Participants sat a pre-test between one and five days prior to the start of each treatment. Because pre-tests before the commencement of new units of work were a common practice at both the primary school and the secondary school, these pre-tests did not alert participants to the forthcoming experiment.

Participants were not given their pre-test scores and did not see their pre-tests again. The pre-test operated to determine the experimental content for each individual participant. Questions on knowledge items which a participant answered correctly in the pre-test resulted in those knowledge items being removed from the experimental content and the post-test score for that participant. The knowledge items referred to in each question which a participant answered incorrectly became the experimental content for that participant for that treatment.

*Retention tests.* Each participant sat his or her own individual retention test one week after the final learning opportunity. Participants who were unavailable for the post test during an 8 to 15 day period following the final experimental lesson were removed from the experimental treatment.

### **Teaching Procedures and Learning Activities**

A number of different activities were developed to provide participants with multiple learning opportunities with the same experimental content. Each different activity provided participants with an additional learning opportunity for each item of content. A number of different activities were created so as to ensure that the learning opportunities did not involve the simple repetition of the to-be-learned material. All the activities provided participants with three-term learning interactions with each knowledge item in the pre-test.

Three activities, the Comprehension Activity, the Creating a Question for the Answer Activity, and the Jumbled Letter Activity, were common to each of the eight experiments which required participants to learn and to remember new knowledge items. The nature of the fourth activity in each experiment varied according to the content of each experiment. In each of the treatments in all eight of the new knowledge item experiments the Comprehension Activity was always the first activity completed by participants. The order of the other activities varied across the eight experiments.

*Learning Activity 1.* The first activity, which will be referred to hereafter as the Comprehension Activity, required participants to answer 12 to 16 comprehension questions about the content of a passage of text provided by the teacher. This passage was issued to each participant at the start of each experimental lesson in each treatment and removed at the end of the lesson. This passage was a critically important document for participants as it contained each of the knowledge propositions tested in the pre and post-tests. Participants could refer to this passage while completing each of the four lesson activities.

*Learning Activity 2.* In the second activity, the Creating a Question for the Answer Activity, participants were given a sheet containing the answers to the comprehension questions from Activity One. Participants were then required to formulate questions that would produce the given answers. For example if the answer to a test question was “It’s the name given to a Martian day” an appropriate question for the answer would be “What is a sol?”

*Learning Activity 3.* A third activity, hereafter referred to as the Jumbled Letter Activity, required participants to fill in the gaps in a passage of teacher-generated text using words and phrases whose letters had been jumbled to form nonsense words which were placed in brackets after each gap. In addition to the jumbled word, a page reference was given to the text for that treatment. Participants could refer to this to help them find the answer. For example in Activity 3 in Experiment 1 the first two sentences in one paragraph read, “Scientists on board our space craft have been talking about a process that turns Mars into an earth-like planet. The process is called \_\_\_\_\_.” In brackets next to the gap was the jumbled phrase (arret formgni p. 4). Participants had to arrange the letters of this word in the correct order to produce *terra forming*.



*Learning Activity 4.* The nature of this activity varied according to the content of the experiment. Descriptions of the fourth activity are provided in the Method section of each experiment.

### **Feedback Procedures for the Learning of New Knowledge Items**

Feedback procedures for each of the lesson activities remained the same across all experimental treatments. The teacher had an important role to play in ensuring that each participant experienced each learning opportunity. During the experimental lesson the teacher moved about the classroom monitoring each participant's progress on the experimental activity. When the teacher was certain that each participant had completed the first activity with the experimental content he called a halt to the class's work and instructed the students to get ready to mark the activity. Next, having made sure that each participant was ready to mark their work, the teacher read out the question and invited participants to respond orally. After a participant's response the teacher indicated whether the answer was correct or not. This ensured that all of the learner interactions with the lesson content consisted of three-term learning interactions, that is, interactions which contained a practice stimulus, a response, and a consequence following that response. After having completed the marking, the participants were instructed to carry on with the task which contained the non-experimental content (Task 2) until the end of the lesson period.

### **Procedures for Controlling the Number of Learning Opportunities**

A number of procedures were employed to control for participants' exposure to the experimental content. First, for each experimental lesson participants worked from hand-outs and work sheets generated by their teacher. At the end of each experimental lesson the hand-outs and all the work sheets were handed back to the teacher. This ensured that the participants had no access to the experimental content outside of the experimental lessons.

For each experimental lesson two instructions were given at the start of the lesson to try to limit the possibility of participants communicating with one another during the lesson. First, participants were instructed to complete both Task 1 and Task 2 without talking. Second, participants were asked to move quietly, again without speaking, to Task 2, the activity involving the non-experimental content, after they had completed Task 1. For each of the experimental lessons the instructions for Task 2 were written on the white board at the front of the class to help participants to move from one activity to the next without talking. Task 2 was specifically designed to ensure that it did not provide any opportunity for the participants to engage with the experimental content from Task 1.

During the experimental lesson, the teacher moved around the room quietly making sure that the participants were following the instructions and not talking. The teacher did not provide any assistance to the participants to answer any of the questions other than encouraging the participant to read the text more carefully to find the answer or by pointing to a section in the text that might provide them with the answer.

For the duration of the experiment the teacher took care to ensure that the homework he set did not stimulate participants to investigate anything concerned with the subject matter of the experimental lessons.

The teacher also took care to ensure that for the week between the final learning opportunity and the post-test the participants did not do any work in class that could provide any opportunity for them to revisit the experimental content.

## EXPERIMENTAL PROCEDURES: PART 2

### **Pilot Study for the Experiments Involving the Learning of New Operations**

The results of a pilot study conducted prior to the commencement of the experiments involving the learning of new mathematics operations demonstrated that it is very important

for the teacher to distinguish between learning how to perform a step in an operation and learning to perform the sequence of steps. In the pilot study a class of Year 12 (Grade 11) mathematics students who did not know how to draw a circle graph was given four learning opportunities to learn the procedure for drawing a circle graph. Seven days after the final learning opportunity the participants were given a post-test. In the post-test only one of the ten participants was able to draw the circle graph correctly. Half of the participants did not even recognize that it was the circle graph that they were required to draw. When the participants were interviewed about their performance in the post test it became clear that many of the participants were not able to perform the sequence of steps required for the procedure because they could not perform one or more of the individual steps of the operation. It became apparent in the pilot study that the participants were being asked to complete two learning tasks at the same time. Not only were participants required to learn how to perform previously unknown component responses but also how to perform the previously unknown sequence of responses required to draw a circle graph. This led the researcher to conclude that a valid test of Nuthall's model with respect to the number of learning opportunities needed to learn and to remember a new operation can only be one in which the students are not required to learn and remember a new sequence of steps until they have first learned how to complete each of the individual steps.

### **Pre-teaching for the Operations Experiments**

Prior to each experiment involving operations, the teacher carried out a task analysis of each 12 to 16 step operation to identify the key component responses and the precise sequence of steps involved in completing each operation. The teacher then developed a schedule for pre-teaching each of the component responses (in a random order) at different times in a single lesson or in separate lessons. Having developed the teaching schedule the teacher then taught each participant how to perform each component response. Participants

were then tested for their ability to perform each of the components (in a random order). Only after each individual had demonstrated that they had learned and remembered how to complete each of the component responses were the participants permitted to sit the pre-test for the operation. Participants who were not able to perform the component steps were withdrawn from the experimental treatment.

### **Testing Procedures for the Operations Experiments**

*Pre-tests.* Participants sat a pre-test for each operation between one and five days prior to the start of each treatment. Participants were accustomed to having pre-tests at the start of each new unit of work in Mathematics and in Art History.

The pre-test determined the experimental content for each participant. Steps for the operation which a participant completed in the correct order in the pre-test were excluded from the number-of-steps-correct measure for that participant. The steps which the participant did not perform in the correct order on the pre-test became the experimental content for that participant for that treatment.

*Retention test.* Each participant sat a retention test one week after the final learning opportunity. This test contained exactly the same questions as the pre-test. Participants who were unavailable for the post test during an 8 to 15 day period following the final experimental lesson were removed from the experimental treatment.

### **Teaching Procedures and Learning Activities**

In Experiments 3 and 10 participants were taught separate mathematical procedures for drawing graphs while for Pilot Study 1 in Chapter 6 participants were taught the procedures for producing art works using different media.

In each experimental lesson for each treatment in Experiments 3 and 10 and for Treatment 1 of Pilot Study 1 participants were required to complete the operation which was the subject for that treatment. A teaching script of instructions was developed for each

experimental lesson for each 12 to 14 step operation for each of the above treatments. This provided the prompts for the to-be-learned operation. The most detailed instructions on how to complete the operation which was the subject of each of the above treatments were provided in the first experimental lesson. In Experimental Lessons 2, 3 and 4, the prompts provided in Experimental Lesson 1 were gradually faded to encourage memorization of the operation. So as not to disadvantage participants who were not able to recall certain steps during the fading process, the detailed instructions which they had received during the first experimental lesson remained available for participants to refer to.

For the first experimental lesson in all the treatments in Experiments 3 and 10 and for the first treatment in Pilot Study 1 the teacher dictated the speed at which the participants completed the operation. The teacher required the participants to complete each step of the operation separately with him to ensure that each participant was able to perform each step correctly. For the remaining lessons in each treatment participants were able to work at their own pace. While the teacher made himself available to assist any participant who encountered difficulties during the experimental lessons, he took care to limit that assistance either to pointing to an instruction or else to asking a participant a question which might help that participant to solve his or her own problem.

For Treatments 2 and 3 of Pilot Study 1 a number of different activities were developed to provide participants with multiple learning opportunities with the same experimental content. Each different activity for Treatment 2 and 3 in Pilot Study 1 provided participants with an additional learning opportunity for each item of content. Different activities were created so as to ensure that the learning opportunities did not involve the simple repetition of the to-be-learned material.

Two activities, the Identifying and Recording the Steps Activity and the Jumbled Steps Activity were common to Treatments 2 and 3 in the Pilot Study. In Treatment 2 the third

learning activity was provided by the Fix the Interview Activity and in Treatment 3 the third learning activity was provided by the Jumbled Steps and Letters Activity. Each of the activities provided a three-term learning opportunity for each sequence of steps in each painting technique operation.

*Learning Activity 1.* The first activity, which will be referred to hereafter as the Identifying and Recording the Steps Activity required participants first to read a passage of text which described the operation for the painting technique (fresco in Treatment 2 and tempera in Treatment 3). The participants then had to identify and record the 13 steps of the operation on an A4 sheet of paper which was numbered 1 to 13 down the left hand margin of the paper.

*Learning Activity 2.* In the second activity, the Jumbled Steps Activity, participants were given a sheet of paper on which the order of the sequence of steps for the painting technique had been deliberately jumbled. Participants simply had to correct the errors in the sequence of steps by writing down the correct order of the 13 steps on an A4 sheet of paper.

*Learning Activity 3.* A third activity, hereafter known as the Fix the Interview Activity, required participants to imagine that they had been provided with the transcript of an actual interview with Michelangelo in which he explained how he had produced the Sistine chapel ceiling fresco. Unfortunately over the years the transcript of the interview had become damaged leaving parts of the transcript unreadable. This activity required participants to fix the transcript of the interview by replacing the unreadable words and phrases with words and phrases that would transform the transcript into a document which provided the correct sequence of steps for Michelangelo's fresco.

*Learning Activity 4.* A fourth activity, hereafter referred to as the Jumbled Step and Letter Activity, was used only for Treatment 3 in Pilot Study 1. In this activity participants were supplied with a sheet containing a sequence of steps that Piero della Francesca took to

produce his tempera painting of *The Baptism*. There were two problems with the sequence of steps. First, the steps had been jumbled up in the wrong order (this was a different order to the jumbled order used in Learning Activity 2) and second, the letters of key words and phrases in each step had been jumbled to form nonsense words. This activity required participants first to arrange the letters of the key words and phrases in the correct order and then to write down the sequence of steps in the correct order.

### **Feedback Procedures**

For the first experimental lesson in each treatment for the three experiments with new operations, participants had to complete each step of the procedure successfully before they were able to move on to the next step. After the completion of each step the teacher provided the correct answer for that step. Only after the teacher was satisfied that each participant had successfully completed the step did the teacher move the participants to the next step. In all subsequent experimental lessons in each treatment the teacher marked each participant's work (after he or she had completed the operation) with the participant sitting alongside the teacher while the other participants continued with another activity.

### **Procedures for Controlling the Number of Learning Opportunities**

The measures taken to control incidental exposure to experimental content and hence to control the number of learning opportunities in the two operations experiments and the pilot study were the same as those used for the eight knowledge item experiments.

## CHAPTER 3

### EXPERIMENT 1

#### EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES FOR UNKNOWN KNOWLEDGE ITEMS IN SETS OF YEAR 7 SCIENCE LESSONS

Following three separate descriptive studies, Nuthall (1993) concluded that students need four exposures to an item of lesson content in order to learn and remember a new knowledge item. Nuthall subsequently altered this model of the learning and remembering process slightly to three to four learning opportunities: “If the student engages with relevant content on three or four distinct occasions (so that at least three or four different representations of content-relevant information come together in the working memory) then a permanent representation of the concept is constructed and transferred to long-term memory” (Nuthall, 1999b, p. 151). In spite of its implications for classroom teaching, the predictive validity of Nuthall’s model has never been tested experimentally.

#### AIM

The aim of Experiment 1 was to see if Nuthall’s finding could be replicated under tightly controlled conditions in which each of the to-be-learned knowledge items were specified in advance. Experiment 1 involved measuring the effects, on the retention of individual Year 7 (Grade 6) students, of four, three and then two learning opportunities, when there is no more than a 2-day inter-study interval between any pair of learning opportunities. A Science topic was chosen for this experiment.

#### METHOD

##### **Participants and Setting**

Participants were students in a mixed ability and mixed gender class of twenty-five Year 7 students at a primary school in Christchurch. Fifteen were boys and ten were girls.



Three of the students had profound learning difficulties which precluded their inclusion in the present experiment. The informed consent procedures described in Chapter 2 were followed for this experiment and the experiment was undertaken in the participants' normal classroom.

### **Lesson Content**

For each 50 minute experimental lesson in each experimental treatment, participants were required to complete two tasks. Task 1 always contained the experimental content and Task 2 always contained the non-experimental content. The content for Task 1 in Experiment 1 came from a self-contained Science unit on astronomy. This unit required participants to imagine that they were involved in a nine month class trip to Mars during which they would be studying Mars and its two giant neighbours, Jupiter and Saturn. Each of the three treatments focused on one of these three planets. The content for Task 2 in each of the three treatments was relevant to the planet being studied but excluded all reference to the content in Task 1. Task 2 typically involved participants in more creative activities such as designing space ships, space costumes and friendly aliens.

### **Testing Procedures**

The procedures for the pretesting and for the retention testing were described in Chapter 2. Only those questions which a participant could not answer on the pre-test were included in the scoring of that student's retention test.

### **Teaching Procedures and Learning Activities**

Descriptions of the teaching procedures for the three learning activities which were used throughout the present experiments (the Comprehension Activity, the Creating a Question for the Answer Activity, and the Jumbled Letter Activity) were given in Chapter 2.

*Learning Activity 1.* The first activity was the Comprehension Activity.

*Learning Activity 2.* The second activity was the Creating a Question for the Answer Activity.

*Learning Activity 3.* For the third activity participants were provided with four pages containing 14 illustrations and diagrams and one page with 14 randomly listed statements about the planet studied in that treatment. Participants had to match the statement to the illustration or diagram and then write the correct statement underneath the illustration or diagram. This activity will be referred to as the Matching Activity.

*Learning Activity 4.* The final activity was the Jumbled Letter Activity.

*Feedback Procedures.* Feedback procedures were described in Chapter 2.

*Controlling the number of learning opportunities.* The number of learning opportunities received by each participant was controlled using the procedures described in Chapter 2.

## **Experimental Treatments**

*Treatment 1.* Experiment 1 consisted of three experimental treatments. Treatment 1 consisted of four experimental lessons on a set of to-be-learned knowledge items on the colonization of Mars. Participants received four opportunities (4 LOs) with each element of the experimental content, one on each day for four successive days. This meant that there was approximately 24 hours between each pair of learning opportunities. The four learning opportunities were provided by the four activities identified above. Task 2 which provided the non-experimental content required participants to design a spaceship to travel to Mars. There were four participants who, due to absence from school, did not experience at least three of the four learning opportunities and because of this they were withdrawn from the treatment.

*Treatment 2.* Treatment 2 consisted of three experimental lessons about the planet Jupiter. The subject matter for these was a set of to-be-learned knowledge items on the planet Jupiter. Participants received three learning opportunities (3 LOs) with the experimental content, one on each day for three successive days. The three participants who did not receive all three learning opportunities were withdrawn from the treatment. Each learning

opportunity was again separated by about 24 hours. The three learning opportunities were provided by the Comprehension Activity, the Question Generation Activity and the Jumbled Letter Activity. The subject matter for Task 2 which contained the non-experimental content was space fashion. Task 2 required participants to create appropriate fashions for astronauts in the 21<sup>st</sup> century.

*Treatment 3.* Treatment 3 consisted of two experimental lessons. The subject matter for these was a set of to-be learned knowledge items about the planet Saturn. All but 4 of the 22 participants received two learning opportunities (2 LOs), one on each day for two successive days. The four who did not receive both learning opportunities were withdrawn from the treatment. As in the previous treatments, the two learning opportunities were separated by about 24 hours. The two learning opportunities were provided by the Comprehension Activity and the Question Generation Activity. The subject matter for Task 2 containing the non-experimental content was aliens. Task 2 required participants to use their imagination to create their own aliens.

## RESULTS

### **Inter-marker Agreement**

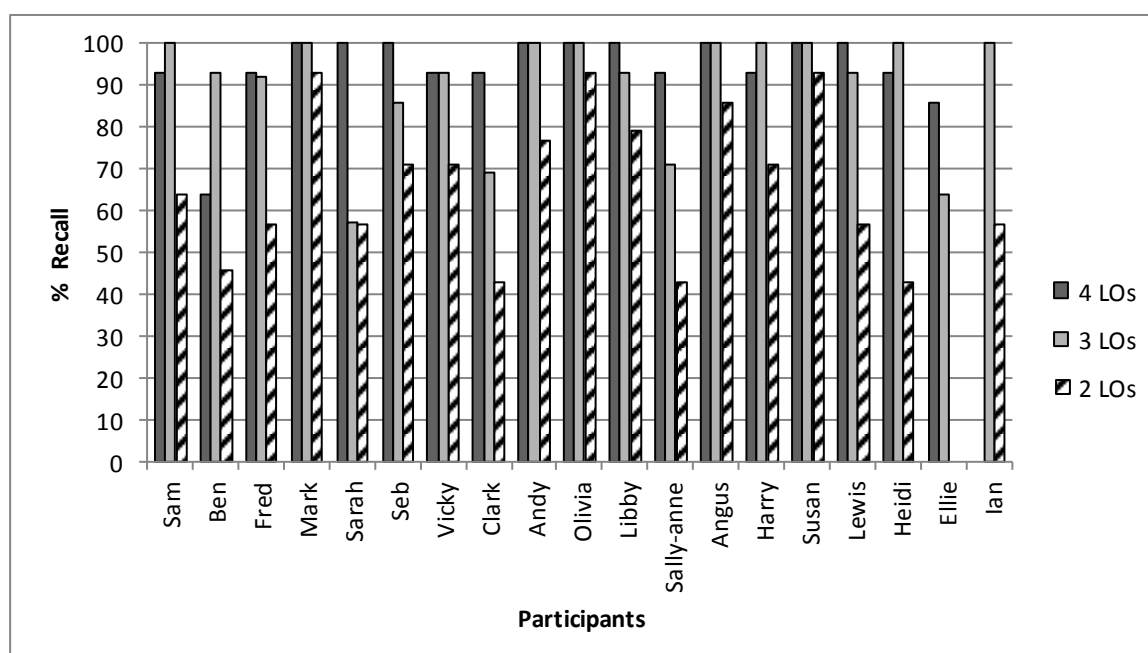
Accuracy checks were conducted by a colleague who check-marked the scoring of all responses in the pre-tests and post-tests. Inter-marker agreement was calculated by using the following formula: “the number of agreements divided by the total of agreements and disagreements multiplied by 100” (Cooper et al., 1987, p. 94). For Experiment 1, inter-marker agreement across all of the tests was 99.3% per cent.

### **Retention of Previously Unknown Knowledge Items**

Questions which a participant answered correctly on their pre-test were not scored on that participant’s individual post-test. For example Libby scored 1/14 in her pre-test for Treatment 1 so the highest score she could achieve was a score of 13/13. Post test scores

(converted to a percentage of previously unknown knowledge items recalled in the post-test) are presented in Figure 2, participant by participant and treatment by treatment.

*Treatment 1.* In Treatment 1, 18 participants received at least three of the four learning opportunities. Their post-test scores ranged from 64 per cent to 100 per cent. Nine participants recalled every answer correctly. Seven participants made just one error and one participant made two errors. Ben (64%) was the only participant who failed to score above 80 per cent in the post-test. His normal class teacher had previously identified Ben as a student who often experienced difficulty settling to new tasks in class.



*Figure 2.* Percentage of unknown knowledge items correctly recalled by each of the participants following each of the three treatments in Experiment 1.

*Treatment 2.* Nineteen participants experienced all three learning opportunities in Treatment 2. The post-test scores ranged from 57 per cent to 100 per cent. Nine participants recalled every answer correctly. All but four participants, Sarah (57%), Ellie (64%), Clark (69%) and Sally-Anne (71%), scored over 80 per cent. In Treatment 2 a number of participants actually performed better in the three learning opportunity treatment than they

had in the four learning opportunity treatment. Ben (64% Treatment 1 and 93% in Treatment 2), Sam (93% Treatment 1 and 100% Treatment 2), Harry (93% Treatment 1 and 100% Treatment 2) and Heidi (93% Treatment 1 and 100% Treatment 2) all performed better in Treatment 2 than in Treatment 1.

*Treatment 3.* Eighteen participants experienced the two learning opportunities in Treatment 3. The post-test scores ranged from 43 per cent (Sally-Anne, Clark and Heidi) to 93 per cent (Olivia, Mark and Susan). Only four participants scored more than 80 per cent in the post-test. For most of the participants there was a marked drop-off in performance from the three learning opportunity treatment to the two learning opportunity treatment. The scores of six participants fell 35 per cent or more in Treatment 3 compared with their scores in Treatment 2. Each participant in Experiment 1 had their poorest performance in Treatment 3 where they experienced only two learning opportunities per unknown knowledge item.

## DISCUSSION

The results from this experiment appear to provide the first experimental confirmation of Nuthall's model of the learning and remembering process, that is, that 11-to-12 year old children require three or four learning opportunities with no more than a two-day inter-study interval between any pair of learning opportunities in order to learn and remember new knowledge items. According to Nuthall's model we would predict that almost all of the 18 participants in Treatment 1 would score more than 80 per cent in their retention tests of previously unknown knowledge items. The fact that all but Ben (64%) did so appears to provide clear support for Nuthall's model. For Treatment 2 Nuthall's model predicts that it was highly likely all 19 participants who experienced all three of the learning opportunities would score more than 80 per cent in the post-test. Fifteen of the 19 children did so. In Treatment 3 where participants received just two learning opportunities Nuthall's model

predicts that for most students their retention would be sharply reduced. This was the case for 14 of the 18 children.

On average, participants performed best in Treatment 1 when most received four learning opportunities and the remainder received three. While some participants performed better in the three learning opportunity treatment than they had in the four learning opportunity treatment, each participant's poorest performance came in the two learning opportunity treatment. The drop-off in retention from the three learning opportunity to the two learning opportunity treatment was marked for many of the participants.

When students receive either three or four learning opportunities to learn a set of new knowledge items, a confirmation of the predictive validity of the Nuthall model occurs when a student gains more than 80 per cent in the post-test and a disconfirmation occurs when a student gains less than 80 per cent in the post-test. Using these definitions Treatments 1 and 2 provided 32 confirmations of the Nuthall model and just 5 disconfirmations.

When students receive just two learning opportunities to learn a set of new knowledge items a confirmation of the Nuthall model occurs when a student scores less than 80 per cent on the post-test while a disconfirmation occurs when a student gains 80 per cent or more in the post-test. With confirmations and disconfirmations defined in this way, Treatment 3 provided a further 14 confirmations and just 4 disconfirmations.

Altogether Experiment 1 provided 46 confirmations and just 9 disconfirmations of the Nuthall model. It might be argued that the content of the Treatment 3 lessons (knowledge about Saturn) was somehow "more difficult" to learn and remember than the content of the Treatment 1 and 2 lessons. Although every effort was made to keep content difficulty and question difficulty the same across the three treatments, this potential confound cannot be completely discounted. Only further replication attempts can answer this question.

## CHAPTER 4

### EXPERIMENT 2

#### EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES FOR UNKNOWN KNOWLEDGE ITEMS IN SETS OF YEAR 13 ART HISTORY LESSONS

Experiment 1 produced the first experimental confirmations and disconfirmations of the Nuthall model of learning and remembering. The number of confirmations (46) heavily outweighed the number of disconfirmations (9). Experiment 2 provided an opportunity to test the Nuthall model with a group of older learners.

The subjects in each of Nuthall's six descriptive studies were primary school students in Year 6 and Year 7. Nuthall suggested in a note to his 1993 report that older students may require fewer learning opportunities to learn and remember new material (Nuthall, 1993). Having demonstrated the predictive validity of the Nuthall model with Year 7 learners in Experiment 1, it seemed important therefore to investigate the generality of the Nuthall model of the learning and remembering process with older learners.

#### AIM

The aim of Experiment 2 was to replicate Experiment 1 with older students, using knowledge items from a higher curriculum level in a different curriculum area. Experiment 2 involved measuring the effects on retention of individual students of four, three and then two learning opportunities, with each pair of learning opportunities separated by an inter-study interval of no more than two days, and with content drawn from the Year 13 (Grade 12) Art History curriculum.

## METHOD

### **Participants and Setting**

Participants were 23 students in a mixed ability class of Year 13 students at a Christchurch secondary school. Seven were boys and 16 were girls. The class was a very compliant group who could be relied upon to work conscientiously at a given task and to adhere to the guidelines set by the teacher. Informed consent was obtained using the procedures described in Chapter 2.

Teaching and testing was conducted in the participants' normal classroom with their normal Art History teacher. During the lessons described below the class was seated in groups containing between three and six students. The classroom was equipped with a data projector.

### **Lesson Content**

As in Experiment 1, each 50 minute experimental lesson in each experimental treatment required participants to complete two tasks with Task 1 always containing the experimental content and Task 2 always containing the non-experimental content. Both the experimental content and non-experimental content came directly from the Year 13 Art History curriculum.

During Experiment 2, participants studied the context, iconography and style of 15th Century Italian Renaissance paintings. Each treatment focused on two paintings, with one providing the experimental content and the other the non-experimental content.

### **Testing Procedures**

The general procedures for the pre-tests were set out in Chapter 2. The pre-tests for this experiment were a twelve to sixteen question tests covering knowledge items on the context, iconography and style of the chosen painting. All questions were constructed response



questions. The participants were supplied with a colour photocopy of the painting for the pre-test. The procedures for the retention tests were described in Chapter 2.

### **Teaching Procedures and Learning Activities**

As in Experiment 1, four very different activities were developed to provide participants with four different experiences with the same lesson content. Again all the activities provided participants with three-term learning interactions with every knowledge item contained in the experimental content. Three of the four activities, the Comprehension Activity, the Question Generation activity and the Jumbled Letter activity, were again used in Experiment 2. Details of these activities are provided in Chapter 2. A new activity had to be created for the Art History unit as the Matching Activity used for the Science unit was not appropriate for Art History. The new activity created for this experiment was the Annotation Activity.

*Learning Activity 1.* The first activity was the Comprehension Activity.

*Learning Activity 2.* The second activity was the Creating a Question for the Answer Activity.

*Learning Activity 3.* The Annotation Activity required participants to annotate a copy of the painting for that treatment, which was provided on an A3 sheet, using a given set of questions and instructions under the headings of Context, Style and Iconography.

*Learning Activity 4.* The final activity was the Jumbled Letter Activity.

*Feedback Procedures.* Feedback on the participants' responses on each lesson activity was provided with the oral answer check procedure described in Chapter 2.

*Controlling the number of learning opportunities.* The number of learning opportunities experienced by each participant with respect to each separate item of content was controlled using the written materials procedure described in Chapter 2.

## Experimental Treatments

*Treatment 1.* Experiment 2 involved three experimental treatments. Treatment 1 consisted of four experimental lessons consisting of a set of knowledge items on the context, iconography and style for the painting *The Misericordia* by Italian artist Piero della Francesca. Participants received four learning opportunities and between each pair of learning activities there was an inter-study interval of approximately 24 hours. The four learning opportunities were provided by the four activities listed above. The non-experimental content for each Task 2 during each of the four experimental lessons was the context, iconography and style for *The Tribute Money* by the Italian artist Masaccio. Three participants did not experience at least three of the four learning opportunities and were withdrawn from the treatment.

*Treatment 2.* Treatment 2 consisted of three experimental lessons about the context, iconography, and style of the portrait of Ginevra de Benci by the Italian artist Leonardo da Vinci. Participants received three learning opportunities with each pair of learning opportunities again being separated by about 24 hours. The three learning opportunities were provided by the Comprehension Activity, the Question Generation Activity and the Annotation Activity. The subject matter for the non-experimental content (Task 2) was the context, style and iconography for Piero della Francesca's *The Baptism*. The eight participants who did not receive the three learning opportunities were withdrawn from the treatment.

*Treatment 3.* Treatment 3 consisted of two experimental lessons. The subject matter for the experimental content was the context, iconography and style for the painting *Lady with an Ermine* by Italian artist Leonardo da Vinci. Participants received two learning opportunities, one on each day for two successive days. The two learning opportunities were provided by the Comprehension Activity and the Question Generation Activity. The subject matter for the

non-experimental part of the lesson was the context, style and iconography for Piero della Francesca's *Victory of Heraclius*. The four participants who did not experience the two learning opportunities were withdrawn from the treatment.

## **RESULTS**

### **Inter-marker Agreement**

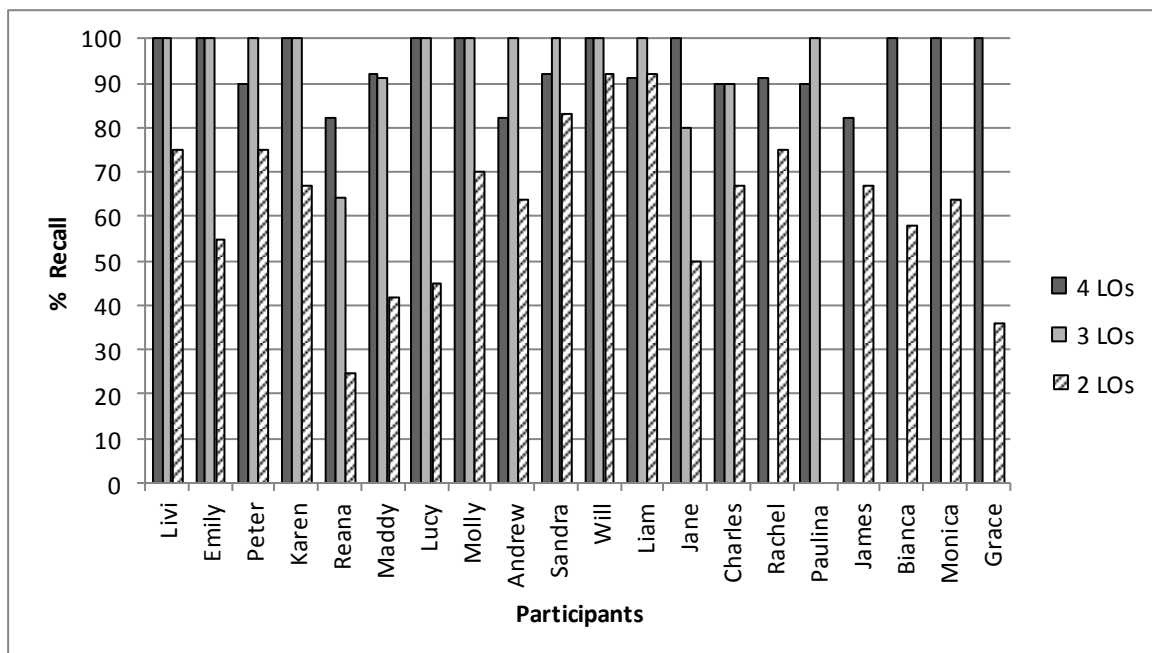
Accuracy checks were conducted by a fellow teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was calculated by dividing the total number of agreements by the total number of agreements and disagreements and then multiplying the result by 100. For Experiment 2 inter-marker agreement was 99.4% per cent.

### **Retention of Previously Unknown Knowledge Items**

Of the 23 participants in the class, 20 completed the pre-tests, the post-tests and the required number of learning opportunities for at least two of the three treatments in Experiment 2. Questions which a participant answered correctly on their pre-test were not scored on that participant's post-test. For example Molly scored 2/12 in her pre-test for Treatment 1 so the highest score she could achieve was a score of 10/10. Post test scores (converted to a percentage of previously unknown knowledge items recalled in the post-test) are presented in Figure 3, participant by participant and treatment by treatment.

*Treatment 1.* In Treatment 1, 17 of the 20 participants received all four learning opportunities with respect to the unknown knowledge items and sat both the pre-test and the post-test. Three participants, Will, Jane and Liam, received just three of the four learning opportunities. Of the 20 participants who received three or four learning opportunities, ten were able to recall the correct responses to all the questions they had been unable to answer on the pre-test, seven were able to recall all but one of the answers to the questions they had not known in the pre-test and three students each made two errors. James and Andrew had the

lowest percentage recall of 82 per cent and thus the range of scores was between 82 to 100 per cent.



*Figure 3.* Percentage of unknown knowledge items correctly recalled by each of the participants following each of the three treatments in Experiment 2.

*Treatment 2.* One of the post-test questions had to be withdrawn from the test because the black and white photocopy of the portrait of Ginevra de Benci issued for the post-test contained a caption below and this caption provided the answer to one of the questions.

In the retention test for Treatment 2, 15 participants experienced all three learning opportunities and sat both the pre-test and post-test. Ten participants retained as high or higher percentage of previously unknown knowledge items as they had following four successive learning opportunities. Only Reana (64%) failed to recall at least 80 per cent of the answers to questions she had been unable to answer on the pre-test. Overall there was little difference between the recall of participants following four learning opportunities with unknown knowledge items and the recall of participants following three learning opportunities with unknown knowledge items.

*Treatment 3.* Nineteen participants experienced both learning opportunities in Treatment 3 and completed both the pre-test and the post test. Recall following two learning opportunities was much more variable. The percentage of knowledge items retained following two learning opportunities ranged from Reana's score of 25 per cent to Liam's and Will's scores of 92 per cent . Only Liam, Will and Sandra (83%) were able to learn and remember 80 per cent or more unknown knowledge items. Apart from Liam all the other participants remembered less following two learning opportunities than they had following three or four learning opportunities.

## DISCUSSION

Experiment 2 provides a systematic replication of Experiment 1 with older students using more difficult knowledge items from a different curriculum area and in a new setting. Overall the results of Experiment 2 closely replicate those of Experiment 1 and appear to provide additional empirical support for Nuthall's claim that, with an inter-study interval of no more than two days between learning opportunities, three or four learning opportunities are needed for students to learn and remember new knowledge items. With a confirmation of the Nuthall model defined as a score of over 80 per cent in the post-test, there were 35 confirmations of the model and one single disconfirmation in Treatments 1 and 2, the four learning opportunity and three learning opportunity treatments respectively. The two learning opportunity treatment produced a further 16 confirmations of the Nuthall model and just 3 disconfirmations. Overall the two experiments which examined the effects on retention of new knowledge items with four, then three and, finally, two learning experiences produced a total of 96 confirmations and just 9 disconfirmations of the Nuthall model.

Nuthall (1996) argued that differences in the post test performance of students in their research studies were not due to differences in ability but rather to differences in the prior knowledge of the students and differences in the way in which students "make use of , and

create, learning opportunities in classrooms” (p.35). The experimental controls employed in Experiment 2 eliminated to a large extent the effects on post-test performance of differences in the participants’ prior knowledge. Any questions which a participant answered correctly in the pre-test were simply excluded from the experimental content for that participant. This left just the participant’s self-directed learning skills as a possible cause of the variability in post-test performance observed following Treatment 3.

This possibility was explored using a series of six short interviews with participants from Treatment 3. Two groups were interviewed, one group which performed well in the retention test for Treatment 3 and a second group which did not. The participants were each asked these two questions: “Were you conscious of doing anything during the activities to help you to remember the material?” and “How were you able to remember the answers for the post-test?”

The three high performing participants were Liam (92%), Will (92%) and Sandra (83%). All three of these participants demonstrated meta-cognitive awareness, that is a self-awareness and self-monitoring of one’s own performance, and all engaged in various forms of self-rehearsal. The three low performing participants were Lucy (45%), Grace (36%), and Maddy (42%). In each case these participants could identify a comprehension strategy but not a remembering strategy. When it came to describing how they recalled the information for the post test their responses were characterised by a large degree of uncertainty.

## CHAPTER 5

### EXPERIMENT 3

#### EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING OPPORTUNITIES DURING THE TEACHING OF A NEW MATHEMATICS PROCEDURE TO YEAR 12 STUDENTS

The results of both Experiment 1 and Experiment 2 provided strong supporting evidence of the Nuthall model's predictive validity for the learning of new knowledge items. Experiment 3 provided an opportunity to explore whether or not the Nuthall model could generalise to another type of learning outcome, the learning of procedures.

Nuthall's model of the learning and remembering process was developed from studies which involved a limited range of learning outcomes. One of the aims of the present set of investigations was to examine the extent to which Nuthall's model could be generalized to different learning outcomes such as the learning of procedures. The *generality* of a model refers to the model's *universality* or *replicability*. A researcher examining the generality of a theory seeks to find an answer to the question: "If I take part or all of the procedures that produced a result and apply them under circumstances that are in some degree different, will I get the same kind of effect?" (Johnston & Pennypacker, 1993, p. 177).

The learning of new procedures played little part in Nuthall's research. In his manual outlining the procedures for identifying the information content of student classroom experiences and predicting student learning, Nuthall pointed out that his team of researchers had yet to determine the "precise combination of experiences necessary for learning procedures...[as] relatively few examples of learning procedures...[had] been identified in our data" (Nuthall, 2001, p. 20). However, later, on the basis of his six descriptive studies, Nuthall claimed that "distinctions among facts, concepts, principles, generalizations, and procedures exist more in the theories of researchers than they do in the minds of the students"

(Nuthall, 2005, p. 911). The claim that his model of the learning and remembering process held true for all types of learning outcomes was also expressed in a footnote in his book *The Hidden Lives of Learners* (Nuthall, 2007, p. 63).

When learning a new spelling word, how to find a word in a dictionary or a new Mathematics skill such as adding fractions, students are required to learn a new procedure, that is “a sequence of responses or operations which must be performed in a particular order if the sequence is to have the desired outcome” (Church, 1999b, p. 42). While some school subjects such as Mathematics are procedure-rich subjects, all subjects require students to learn procedures of varying degrees of complexity. Some researchers argue that the way in which new procedures are taught plays a critical role in determining whether or not the student will learn and remember a new procedure (Church, 1999b; Sulzer-Azaroff & Mayer, 1991).

It will be recalled from the discussion in Chapter 2 of a pilot study conducted with ten Year 12 (Grade 11) Mathematics students that the failure of all but one of the students to learn and remember the procedure for drawing a circle graph was likely to have been due to the way in which the students had been taught the procedure. In the pilot study students had been provided with four learning opportunities, one each day on four successive days, to learn not only how to perform each of the 12 steps of the procedure but also how to perform the sequence of steps required to complete the procedure. This led the researcher to conclude that a valid initial test of the Nuthall model needed to be a test of just the participant’s ability to recall the sequence of steps of which the procedure is comprised.

### AIM

The aim of the Experiment 3 was to see if Nuthall and Alton-Lee’s finding with respect to the number of learning opportunities required to learn and remember a set of previously unknown knowledge items could also be applied when learners were faced with learning the



sequence of steps necessary in order to complete an unknown procedure. Experiment 3 involved measuring the effects on the retention of individual students of four, three and then two learning opportunities with no more than a two day inter-study interval between any pair of learning opportunities, with respect to the learning of three separate procedures presented during sets of Year 12 (Grade 11) Mathematics lessons.

## METHOD

### **Participants and Setting**

The participants for the four experimental conditions in Experiment 3 were a class of ten, average to low ability, mixed gender Year 12 Mathematics students at a secondary school in Christchurch. Members of this class had been placed in an alternative Mathematics programme in Year 11 because their performances in Year 10 suggested they would not be able to manage the full Year 11 National Certificate of Education Achievement (NCEA) course. In Year 12 these students had opted to continue with the less challenging alternative Mathematics programme which provided them with the opportunity to gain internally assessed Unit Standard credits. Several of the students in this class had been diagnosed with specific learning difficulties. Many of the problems the students encountered in Mathematics stemmed from the fact that they had poor basic Mathematics skills. Despite the fact that the students found Mathematics challenging they had very sound attitudes and a willingness to learn. The number of participants who entered each treatment and the number of participants who completed each treatment are given in the Results section below. The informed consent of these students was obtained using the Informed Consent procedures described in Chapter 2. Teaching and testing was conducted in the participants' normal Mathematics classroom with their usual classroom teacher.

## **Lesson Content**

For each 50 minute experimental lesson in each experimental treatment, participants were required to complete several tasks. Task 1 always consisted of a warm up activity aimed at developing participants' fluency in basic multiplication facts and in the addition and subtraction of algebraic terms. Task 2 always contained the experimental content. The content for Task 2 came from a self-contained unit on drawing graphs from the Year 12 Alternative Mathematics curriculum. This experiment required participants to learn three 12 to 14 step procedures for graphing a parabola, a circle and a line. Each of the three experimental treatments focused on one of the three graphs above. Task 3 in each of the treatments required participants to complete a set of exercises on either a Statistics or a Trigonometry unit.

## **Pre-teaching**

The teacher carried out a task analysis of each 12 to 14 step procedure to identify the key component responses and the precise sequence of steps which go up to make each procedure. The teacher then developed a schedule for teaching each of the component responses at different times in a single lesson or in completely separate lessons and in a random order to prevent participants, as far as possible, from intuitively working out the correct sequence of steps needed to complete the procedure for that treatment. Having developed the teaching schedule the teacher then taught each of the component responses to each participant. Participants were then tested for their knowledge of each of the component responses. Only after each individual had demonstrated that they had learned and remembered how to complete each of the key component responses of the procedure were the participants permitted to sit the pre-test for the procedure.

## Testing Procedures

*Pre-tests.* The general pretesting procedures are described in the Chapter 2. For the pre-test for each procedure in Experiment 9 participants were issued with an A3 sheet of paper which was divided into two. At the top of the page on the left-hand side of the A3 sheet was an equation for the graph which was the subject of that treatment. Below the equation was some graph paper on which the participant could attempt to draw the graph represented by the equation on the  $x$  and  $y$  axes already drawn on the graph paper. The right-hand side of the A3 sheet had been left clear to enable the participant to carry out any working he or she considered necessary. The pre-test determined the experimental content for each participant. Steps for the procedure which a participant completed correctly in the right order in the pre-test were excluded from the experimental content for that participant. The steps which the participant could not perform in the correct order in the pre-test became the experimental content for that participant for that treatment.

*Retention tests.* The retention test was the same as the pre-test. The retention tests were administered 7 days after the final lesson in each treatment using the procedures described in Chapter 2.

## Teaching Procedures and Learning Activities

In each experimental lesson for each treatment in Experiment 3, participants were required to complete the operation which was the teaching aim for that treatment. A teaching script was developed for each experimental lesson for each 12 to 14 step procedure. This provided the prompts for the to-be-learned procedure. The instructions were written on the left-hand side of an A3 sheet. On the right-hand side of the sheet the participants were provided with graph paper. A set of  $x$  and  $y$  axes had been drawn on the graph paper to assist the participants in drawing their graphs. The most detailed instructions were provided in the first experimental lesson. In Lessons 2 to 4 the prompts provided in the first experimental

lesson 1 were gradually faded to encourage memorization of the procedure. In each successive experimental lesson after the first, participants received a new script of instructions which was less detailed than the instructions provided in the previous experimental lesson. However, participants were always issued with the instructions they had received from all the previous experimental lessons in that specific treatment so that each participant was always able to refer back to a previous script if he or she needed to.

The teacher began the task with the experimental content by explaining to the class that they were going to learn a procedure for drawing a specific type of graph. Having identified the graph, the teacher explained to the class that the procedure for drawing the graph had been broken down into a series of steps on the instruction sheet on the left-hand side of their A3 sheet and that these instructions would enable them to draw the graph on the graph paper supplied on the right-hand side of the sheet. For the first experimental lesson in all three treatments in Experiment 3 the teacher dictated the speed at which the participants completed the procedure. The teacher required the participants to complete each step of the procedure separately with him to ensure that each student was able to perform each step in the correct order. For the remaining lessons in each treatment participants were able to work at their own pace. While the teacher made himself available to assist any participant who encountered difficulties during the experimental lessons, he took care to limit that assistance either to pointing to an instruction or else to asking a participant a question which might help that participant to solve his or her own problem.

Feedback following participant responses was provided using the feedback procedures described in Chapter 2 and the number of learning opportunities was controlled by the use of written task instructions which were always collected at the end of each lesson so that no additional practice could occur outside these lessons.

## Experimental Treatments

*Treatment 1.* Experiment 3 consisted of four experimental treatments. Treatment 1 consisted of four experimental lessons on how to draw a parabola graph. In the two week period prior to the beginning of the experimental lessons, participants were taught how to perform each of the component responses of which the procedure was comprised. The teacher took care to teach the component responses in a random order both at different times during a lesson and in separate lessons. The procedure for drawing a parabola graph was broken down into its eleven key components. These were being able to: (a) recognize the formula for a parabola; (b) recognize the shape of a parabola graph; (c) locate specified co-ordinates on a graph; (d) identify the  $x$  and  $y$  axes on a graph; (e) find the  $x$  intercepts from the equation for a parabola; (f) find the mid-point between the two  $x$  intercepts; (g) find the mirror line between two  $x$  intercepts; (g) find the  $y$  intercept from the equation for a parabola; (h) add integers; (i) multiply integers; and, finally, (j) explain the concept of symmetry. Participants had to demonstrate that they were able to complete each of the component skills before they were permitted to sit the pre-test. Participants who were able to complete the procedure in the pre-test were withdrawn from Treatment 1. Two participants were able to complete the procedure in the pre-test and were withdrawn from the treatment. One participant failed to meet the requirement to experience at least three of the four learning opportunities and was also withdrawn from the treatment.

Participants received four opportunities, one on each day for four successive days, to complete the 14 step procedure for drawing a parabola graph. This meant that approximately 24 hours elapsed between each pair of learning opportunities. Task 3 which provided the non-experimental content required participants to complete a set of exercises for a Statistics unit.

The 14 step procedure for drawing the parabola  $(x + 4)(x - 3)$  was as follows:

1. Recognise it is a parabola graph

2. Find the first  $x$  intercept by making the bracket  $(x + 4) = 0$ , i.e.  $x = -4$
3. Mark this point on the  $x$  axis
4. Find the second  $x$  intercept by making the bracket  $(x - 3) = 0$ , i.e.  $x = 3$
5. Mark this on the  $x$  axis
6. Find half way between the two intercepts by first counting the squares between the two intercepts, i.e. 7 squares
7. Divide this number by 2, e.g.  $7/2 = 3.5$
8. Travel 3.5 squares from any one of the  $x$  intercepts towards the other and mark it with an X to identify the mid-point
9. Write in the co-ordinates for the mid-point next to X, e.g.  $(-0.5, 0)$
10. Draw a vertical line through the mid-point to form a line of symmetry for the parabola
11. Find the  $y$  intercept by first making  $x = 0$  in both brackets, e.g.  $y = (0 + 4)(0 - 3)$
12. Multiply  $4 \times -3 = -12$
13. Mark in the  $y$  intercept  $(0, -12)$  with an X
14. Draw in a *symmetrical* parabola which passes through both  $x$  intercepts and the  $y$  intercept.

In the third of the four experimental lessons Fergus asked for a new A3 sheet after he realized that he had made an error in drawing his graph. The teacher made the decision to provide an additional A3 sheet to allow Fergus to make his correction as this was what would have happened in a normal classroom lesson.

*Treatment 2.* Treatment 2 consisted of three experimental lessons on the 12 step procedure for drawing a circle graph. Teaching of the procedure for drawing the circle graph did not begin until the completion of Treatment 1 to prevent confusion in the minds of participants who, while they had acquired the knowledge to complete the procedure for drawing the parabola, had not had the opportunity to master the procedure. As in Treatment 1, in the two week period prior to the beginning of the experimental lessons participants were taught the prerequisite skills and knowledge required to complete each of the key components of the procedure. Care was taken to teach the component skills and knowledge in a random

order and in different lessons so as to prevent the possibility of participants working out the correct sequence of steps needed to draw the circle graph for themselves. Again participants had to demonstrate that they were able to complete each of the component skills before they were permitted to sit the pre-test. None of the nine participants who sat the pre-test were able to complete the procedure after being taught the component skills and knowledge and hence all nine were eligible for the treatment. One student did not experience the three learning opportunities and was withdrawn from the treatment.

The 12 step procedure for drawing a circle graph was broken down into its five key components. These were being able to: (a) recognize the formula for a circle; (b) find a point in all four quadrants of a graph using a set of co-ordinates; (c) find the square root of 9, 16, 25, 49, 64 and 81; (d) find the  $x$  and  $y$  co-ordinates from any given circle graph formula; and, (e) locate points on a graph for a given number of squares above, below, to the right or to the left of a given point on a graph.

The 12 step procedure for drawing the circle  $(x + 2)^2 + (y - 1)^2 = 9$  was as follows:

1. Recognise the need to draw a circle graph from the equation  $(x + 2)^2 + (y - 1)^2 = 9$
2. Find the square root of 9
3. Write down what the radius is equal to.
4. Find the  $x$  co-ordinate by making the bracket  $(x + 2) = 0$ , e.g.  $x = -2$
5. Find the  $y$  co-ordinate by making the bracket  $(y - 1) = 0$ , e.g.  $y = 1$
6. Locate the point  $(-2, 1)$  on the graph and mark it with an O.
7. Next to the O write down the co-ordinates for the centre of the circle
8. From O go right 3 squares and mark it with an X
9. From O go up 3 squares and mark it with an X
10. From O go left 3 squares and mark it with an X
11. From O go down 3 squares and mark it with an X.
12. Draw the circle graph by drawing a circle through each of the 4 points marked with an X.

Participants received three learning opportunities, one each day for three successive days, to complete all 12 steps of the procedure for drawing a circle graph. The non-experimental content (Task 3) for the lesson required participants to complete a set of exercises for a Trigonometry unit.

*Treatment 3.* Treatment 3 consisted of two experimental lessons on two successive days on how to draw a line graph. Participants again had to demonstrate that they possessed the necessary prerequisite knowledge and skills to complete each step of the procedure for drawing a line graph from a linear equation before they were eligible to sit the pre-test. Participants were taught the component skills in the two week period prior to the start of the experimental lessons. The teacher again took care to teach the component responses in a random order both at different times during a lesson and in separate lessons to try to prevent participants from learning the sequence of steps which would enable them to draw the line graph. One participant was able to complete the procedure at the pre-test and was withdrawn from the treatment. Another participant missed the pre-test and did not receive both learning opportunities and so was also withdrawn from the treatment.

The 13 step procedure for drawing the line graph was broken down into its six key components. These were being able to: (a) recognize  $y = x$  as the most basic formula for a line graph; (b) find the highest common factor with which to divide a linear equation such as  $3y = 12x + 18$  so that the equation can be expressed in the form  $y = mx + c$ ; (c) find the value of  $y$  in a linear equation by making  $x = 0$  and working out the answer; (d) find the value of  $x$  in a linear equation by making  $y = 0$ ; (e) changing the subjects of algebraic equations; (f) working out the answer by making  $x$  the subject.

The 13 step procedure for drawing the line graph  $3y = 12x + 18$  was as follows:

1. Recognise that  $3y = 12x + 18$  is a line graph
2. Divide the equation by 3 ( $y = 4x + 6$ )



3. Find  $y$  by making  $x = 0$ ;  $y = 4 \times 0 + 6$
4. Solve this;  $y = 6$
5. Write down the co-ordinates for the first point (0, 6)
6. Mark the point on the graph
7. Find  $x$  by making  $y = 0$ ; e.g.  $0 = 4x + 6$
8. Make  $x$  the subject by subtracting 6 from both sides  $-6 = 4x$
9. Divide both sides of the equation by 4 ( $-6/4 = m$ )
10. Solve this;  $x = -1.5$
11. Write the co-ordinates for the second point (-1.5, 0)
12. Mark the point on the graph
13. Draw in the line between the 2 points

Participants received two learning opportunities on successive days. After completing the activities containing the experimental content in each of the two experimental lessons the participants continued their work on their Trigonometry unit.

Because Treatment 3 produced an unexpected result it was replicated using the same participants and a fourth procedure, the drawing of a hyperbola graph.

*Treatment 4.* In the two week period prior to the pre-test participants were taught the component skills for drawing a hyperbola in ten minute blocks sometimes at the beginning of and sometimes during each Mathematics lesson. The teacher took care to teach the component responses in a random order and in separate lessons to try to prevent participants from learning the sequence of steps which would enable them to draw the hyperbola graph.

Two participants were withdrawn from the treatment because they were absent during the period of the experimental treatment.

The 13 step procedure for drawing a hyperbola graph was broken down into its eleven key components. These were being able to (a) recognise  $xy = 6$  as the basic formula for a hyperbola graph ; (b) locate a given co-ordinate; (c) sketch both positive and negative vertical lines such as  $x = 3$  and  $x = -3$ ; (d) sketch both positive and negative horizontal lines such as

$y = 2$  and  $y = -2$ ; (e) explain what an asymptote is; (f) sketch a vertical asymptote such as  $x = -2$ ; (g) sketch a horizontal asymptote such as  $y = 3$ ; (h) produce the positive factors of a given positive number on an  $x/y$  factor grid; (i) find the negative factors of a given positive number on an  $x/y$  factor grid; (j) find the new origin of a graph at the intersection of the vertical and horizontal asymptotes; (k) plot co-ordinates when the origin of a graph has been shifted to the intersection of the two asymptotes. None of the eight participants who sat the pre-test were able to complete the procedure after being taught the component skills and hence all eight were eligible for the treatment.

The 13 step procedure for sketching the graph of the hyperbola  $y = \frac{4}{x+1} - 2$  was as follows:

1. Recognise  $y = \frac{4}{x+1} - 2$  as a hyperbola graph
2. Work out the vertical asymptote by making  $(x + 1) = 0$ , i.e.  $x = -1$
3. Draw in the vertical asymptote at  $x = -1$
4. Work out the horizontal asymptote, i.e.  $y = -2$
5. Draw in the horizontal asymptote at  $y = -2$
6. Mark the new Origin with an O at the intersection of the vertical and horizontal asymptotes (i.e. at  $(-1, -2)$ )
7. Draw the  $x/y$  factor grid
8. Write in the positive and negative factors of 4 in the  $x/y$  factor grid
9. Plot one positive co-ordinate for the graph.
10. Plot a second positive co-ordinate for the graph
11. Plot one negative co-ordinate for the graph
12. Plot a second negative co-ordinate for the graph
13. Sketch in the graph by linking the two positive co-ordinates with a curved line and then by linking the two negative co-ordinates with a curved line.

Participants received two learning opportunities, one on each day for two successive days to complete the 13 step procedure for drawing a hyperbola graph. This meant that there was approximately 24 hours between the two learning opportunities.

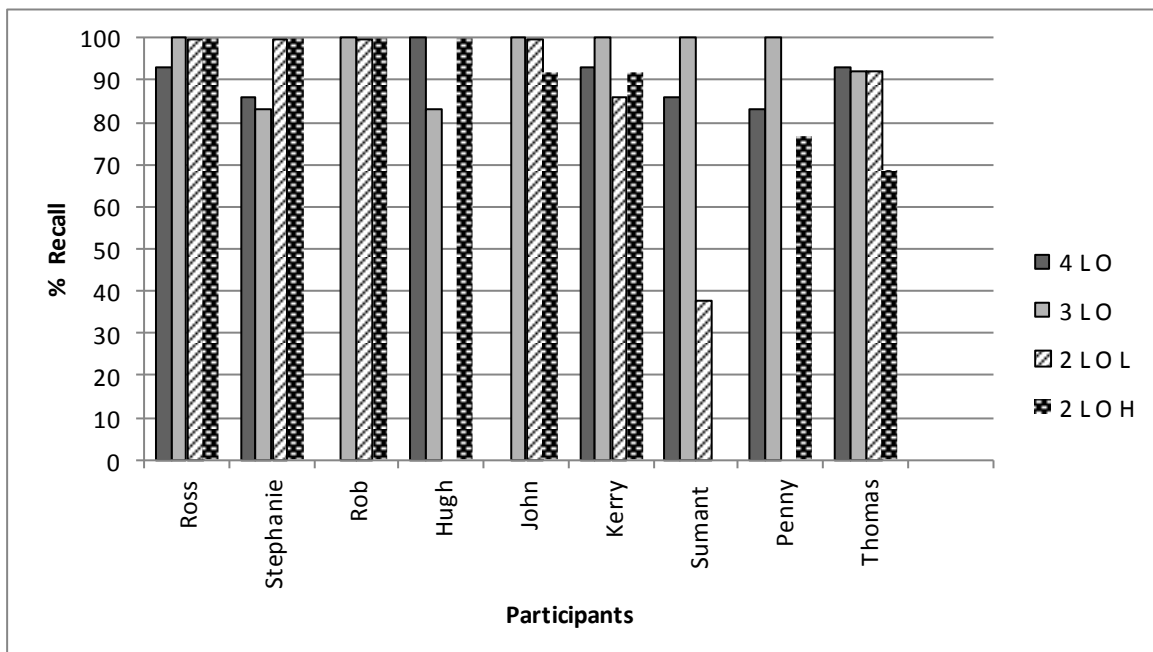
## RESULTS

### Inter-marker Agreement

Accuracy checks were conducted by a fellow teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of agreements and disagreements, and then multiplying the result by 100. For Experiment 3, inter-marker agreement was 97.6% per cent.

### Retention of Previously Unknown Procedures

Steps for the procedure which a participant completed correctly in the right order in their pre-test were not scored on the participant's retention test. The retention test scores for each participant in each of the four treatments are presented in Figure 4.



*Figure 4.* Percentage of steps performed correctly by each of the participants following each of the four treatments (parabola - 4 learning opportunities (4 LO), circle graph – 3 learning opportunities (3 LO), line graph – 2 learning opportunities (2 LO L), and hyperbola graph – 2 learning opportunities (2 LO H)) in Experiment 3.

*Treatment 1.* Ten participants sat the pre-test after demonstrating their understanding of the component skills and knowledge in a short test. Two of the ten participants, John and Rob, were able to complete the graph in the pre-test which meant they were excluded from the treatment. Absences from two of the four experimental lessons ruled out a third member of the class, Gavin, from participation in the treatment. His frequent absences at critical times during the course of the experiment meant that he had to be withdrawn from the experiment altogether. Six of the seven remaining participants received four learning opportunities on four successive days while Sumant, due to his absence in the first experimental lesson, received two learning opportunities on one day and then a further two learning opportunities on two successive days. As can be seen from Figure 3, Hugh was the only participant to complete each step in the procedure for drawing a parabola without making any errors. Ross, Stephanie, Kerry, Penny and Thomas completed the procedure but each made a single error. Ross, Stephanie and Kerry failed to produce a sufficiently symmetrical parabola, Penny failed to mark in the y intercept correctly and Thomas wrote the order of his co-ordinates down incorrectly. Sumant made two errors; he failed to draw a symmetrical graph and he failed to write in the co-ordinates of the mid-point.

*Treatment 2.* Nine participants sat the pre-test for the circle graph after having demonstrated in a short test that they could perform the component skills. Even though each of the participants had been able to identify the equation that produced a circle graph in the pre-test only one of the participants, Rob, was able to produce a circle graph but even his effort bore little resemblance to the correct graph. All nine of the participants experienced the three learning opportunities. Kerry missed the second learning opportunity and so in the third experimental lesson she received the second and third learning opportunities. All the other participants received one learning opportunity each day for three successive days. In the post-test six of the nine participants who received three learning opportunities completed the

procedure for drawing a circle without making a single error. Thomas forgot to give the co-ordinates of the centre of the circle while Hugh and Stephanie both made two errors having failed to find the radius correctly.

*Treatment 3.* Eight participants sat the pre-test after they had first demonstrated their understanding of the prerequisite skills and knowledge in a short test. Only Hugh did not sit the pre-test. He was absent on a school trip during the period of this experimental treatment. Penny succeeded in completing the procedure for the line graph in the pre-test and, as a consequence, was withdrawn from the experimental treatment. Four participants failed to complete a single step in the pre-test. Of the remaining three participants, John completed seven steps, Kerry completed five and Rob completed four. In the retention test Ross, Rob, John and Stephanie completed all of the steps of the procedure correctly and Kerry and Thomas each made one error. Kerry made a simple division error in Step 2 of the procedure and Thomas failed to give the co-ordinates of the second point in Step 11. Only Sumant failed to carry out the procedure effectively; he was able to complete only five of the steps correctly. An analysis of his retention test revealed that he had forgotten how to perform key component skills and this prevented him from being able to complete the procedure.

*Treatment 4.* Eight participants qualified to sit the pre-test after having demonstrated that they had learned the component skills in a short test. Sumant was withdrawn from the treatment because he was absent for both the pre-test and the two experimental lessons. In the pre-test none of the eight participants was able to complete a single step in the procedure. In the retention test, four participants, Rob, Hugh, Ross, and Stephanie, completed all thirteen steps of the procedure correctly while two students, Kerry and John, each made the same single error; each failed to draw in the vertical asymptote correctly. Kerry and John had remembered the correct procedure but in each case their inability to perform one of the component skills correctly prevented them from performing the procedure with complete

accuracy. Penny completed all but three steps of the procedure; she failed to plot the two negative co-ordinates and, as a result, did not complete the negative arm of the hyperbola which should have been in the third quadrant of the graph. Thomas' performance was the worst of the group. He was not able to locate the right position for either the  $x$  asymptote or the  $y$  asymptote and so did not find the correct location for the new origin. However, he did manage to plot the two positive and two negative co-ordinates correctly from the point he believed was the new origin. At the time of the post-test Thomas' health was not at its best. He suffered from an on-going ailment which had flared up in the week of the retention test causing him much discomfort. In addition he was heavily involved in out-of-class school activities and, as a result, had, by his own admission, lost some of his motivation for school work.

## DISCUSSION

The results of both Treatment 1 and Treatment 2, the four and three learning opportunity treatments respectively, provided supporting evidence of the predictive validity of the Nuthall model with respect to the number of learning opportunities required to learn and remember a new procedure. Treatments 1 and 2 produced not a single disconfirmation of the model and a total of 16 confirmations of the model.

Treatments 3 and 4, both two learning opportunity treatments, provided 12 disconfirmations of the Nuthall model with respect to the number of learning opportunities required to learn a new procedure. The results of Treatments 3 and 4 suggest that students may require as few as two learning opportunities to learn and remember the sequence of steps required to perform a Mathematics procedure, provided they already know how to perform the component steps of the procedure. This is a major finding for it suggests that the Nuthall

model may not apply to all types of learning outcomes and therefore it may not be as generalisable as Nuthall has claimed.

For the Mathematics teacher in a real-life classroom the results of this experiment clearly demonstrate that the way a new Mathematics procedure is taught can play a crucial role in determining whether or not a student will learn and remember that new procedure.

CHAPTER 6  
PILOT STUDY 1  
EFFECTS ON RETENTION OF REDUCING THE NUMBER OF LEARNING  
OPPORTUNITIES FOR UNKNOWN PROCEDURES IN SETS OF YEAR 13 ART  
HISTORY LESSONS

It will be recalled that the two 2-learning opportunity treatments, Treatments 3 and 4, in Experiment 3 unexpectedly produced a larger number of disconfirmations (12) than confirmations (3). Six out of 7 Year 12 participants in Treatment 3 and 7 out of 8 Year 12 participants in Treatment 4 were able to graph an equation in a post-test one week after receiving the second of only two learning opportunities. This suggested that further experiments investigating the number of learning opportunities which students need in order to learn and remember new procedures should be undertaken. To this end a further experiment investigating the learning and remembering of procedures was conducted using different content and different students.

AIM

The aim of Pilot Study 1 was to see if the findings from Experiment 3 with respect to the number of learning opportunities required to learn and to remember three to-be-learned Mathematics procedures could be replicated during the teaching of three Art History procedures. Pilot Study 1 involved measuring the effects on the retention of individual students of four, three and then two learning opportunities (with inter-study intervals of two days or less) during the teaching of painting procedures included as part of a set of Year 13 Art History lessons.



## METHOD

### **Participants and Setting**

The participants for the three experimental conditions in Pilot Study 1 were from a class of mixed ability and mixed gender Year 13 Art History students at a secondary school in Christchurch. This Year 13 (Grade 12) Art History class consisted of five boys and ten girls with an average age of 17 years. The number of participants who entered each treatment and the number of students who completed each treatment are given in the Results section below. The informed consent procedures described in Chapter 2 were followed for this experiment and the teaching and testing was conducted in the participants' normal classroom with their usual Art History teacher.

### **Lesson Content**

Participants were required to complete two main tasks in each of their 50 minute Art History lessons. The content for both tasks came directly from the Year 13 Art History curriculum. The first task in each lesson contained the pilot study content. The pilot study content for all three treatments in Pilot Study 1 came from a self-contained unit on Renaissance painting techniques. The second task (the non-pilot-study task) in Treatment 1 required participants to complete a series of activities on the impact of the Florentine economy on art works. The second task in Treatment 2 required students to complete a series of questions on the impact of scientific developments on art works in Florence while the second task in Treatment 3 required participants to investigate the style, context and iconography of Leonardo Da Vinci's portrait of Mona Lisa.

### **Testing Procedures**

*Pre-tests.* The general pretesting procedures are outlined in Chapter 2. The pre-test for each treatment in Pilot Study 1 asked participants to perform or describe a 13 step procedure. Participants sat the pre-test only after they had demonstrated that they had learned and

remembered how to complete each of the essential component responses for the procedure. For Treatment 1 the participants received an A3 sheet of paper with a rectangle measuring 20cm by 16cm which represented the wall on which an artist would draw up a single-point perspective grid. In the top right-hand corner of the A3 sheet were instructions to the participant to set out an Albertian single-point perspective system as accurately as possible. For the pre-test in Treatment 2 participants were required to write down in the correct order the steps Michelangelo would have taken to produce his fresco on the ceiling of the Sistine Chapel. For the pre-test in the third treatment participants were asked to write down the sequence of steps that Piero della Francesca would have taken to produce the tempera painting, *The Baptism*. For all three treatments the pre-test determined the experimental content for each participant. Steps for each procedure which a participant completed correctly in the pre-test were excluded from the experimental content for that participant. The steps which the participant did not know how to perform in the pre-test became the experimental content for that participant for that treatment.

*Retention tests.* The retention tests were the same as the pre-tests and were administered 7 days after the final lesson in each treatment using the procedures described in Chapter 2.

### **Teaching Procedures and Learning Activities**

*Pre-teaching of the prerequisite component responses.* The teacher carried out a task analysis for each 13 step procedure to identify the key component responses and the precise sequence of steps required for each procedure. The teacher's first task was to teach each participant the prerequisite skills and knowledge which would enable them to complete each step of the procedure. The challenge for the teacher was to teach the prerequisite skills and knowledge in such a way as to prevent participants from working out for themselves the sequence of steps needed to complete the procedure for that treatment. In an attempt to preclude this as far as possible, the teacher taught the prerequisite skills and knowledge in

isolation in a random order in a number of separate lessons which were often widely spaced. Participants were then tested for their knowledge of each of the components. Only after each individual had demonstrated that they had learned and remembered how to complete each of the key component responses of the procedure were the participants permitted to sit the pre-test for the procedure.

In each pilot study lesson for each treatment in Pilot study 1 participants were required to complete the operation which was the teaching objective for that treatment. A teaching script of instructions was then developed for each learning opportunity in each pilot study lesson for each procedure. This provided the prompts for the to-be-learned procedure. More detailed scripts of instructions were provided for the first two pilot study lessons of each treatment. In an effort to encourage each participant to rely less on the instructions and more on their memories, less detailed instructions were provided for the third learning opportunity in Treatment 1 and Treatment 2 and even less detailed instructions were provided for the fourth learning opportunity in Treatment 1. To avoid a situation in which a participant could not complete the learning opportunity because they had forgotten a crucial step and the less detailed instructions failed to prompt the forgotten step, participants received in each new pilot study lesson not only their new instructions but also the instructions they had received from their previous pilot study lessons. The scripts containing the instructions for completing each procedure had been trialled with an adult learner to ensure that the instructions were clear and, when followed correctly, produced the desired outcome. Feedback was provided using the procedures described in Chapter 2 and the number of learning opportunities was controlled by using written material which students could only access during the lesson (as described in Chapter 2).

Each learning activity with the pilot study content in each of the four pilot study lessons in Treatment 1 required participants to complete the sequence of steps needed to construct an

Albertian single point perspective grid. To add some variety to these learning activities participants were given a different set of measurements with which to construct their Albertian single point perspective grid in each of the four learning activities in Treatment 1.

Descriptions of the teaching procedures for the learning activities for the content of Treatment 2 and Treatment 3 (the Identifying and Recording Steps Activity, the Jumbled Steps Activity, the Fix the Interview Activity, and the Jumbled Steps and Letters Activity) are given in Chapter 2.

*Learning Activity 1.* The first activity which was used in both Treatments 2 and 3 was the Identifying and Recording Activity.

*Learning Activity 2.* The second activity was the Jumbled Steps Activity. This activity was also used in both Treatments 2 and 3.

*Learning Activity 3.* The third activity was the Fix the Interview Activity. This activity was only used in Treatment 2.

*Learning Activity 4.* The fourth activity was the Jumbled Steps and Letter Activity. This activity was only used in Treatment 3. (In this activity the steps had been jumbled in a different order to that used in Learning Activity 2, the Jumbled Steps Activity.)

### **Pilot Study Treatments**

*Treatment 1.* In Treatment 1 participants were given four lessons to learn the 13-step procedure for drawing an Albertian single point perspective grid. Prior to the start of the experimental lessons the teacher taught the participants the necessary component skills and knowledge required to successfully complete the procedure. This meant that participants had to be able to define the following concepts: *orthogonals* (diagonal lines that can be drawn along receding parallel lines to the vanishing point); *transversals* (horizontal lines showing depth in a painting); *viewing distance* (the distance between the viewer and the art work); *ideal viewing point* (the ideal spot from which to view the art work); *vanishing point* (the

centre point on the horizon line); *braccia* (the unit of measurement that was used to determine the distance between the orthogonals at the base of the fresco; one braccia was equivalent to one third of the height of a man); *fresco* (a wall painting); and *horizon line*. Participants also had to demonstrate the ability to identify each concept from a series of simple diagrams which illustrated each concept individually. Only those participants who demonstrated that they had learned the prerequisite skills and knowledge went on to sit the pre-test. In Treatment 1 participants who met the criteria for participation in the treatment received four learning opportunities to complete all 13 steps of the procedure for creating an Albertian single-point perspective system for a fresco. The timetabling of the Year 13 Art History classes meant that participants received one learning opportunity each day for three successive days and a fourth learning opportunity two days later. This meant that there were approximately 24 hours between each of the first three learning opportunities and a 48 hour interval between the third and fourth learning opportunities. Task 2, which provided the non-experimental content, required participants in each of the four experimental lessons to complete a set of exercises on the economy of fifteenth century Florence in preparation for a test.

The 13 step procedure for creating a single point perspective system for a fresco was as follows:

1. Draw a man (Stick Man I) to the right of the fresco to a specified height (Participants used an A3 sheet of paper in lieu of the wall)
2. Draw a horizon line at the height of the stick man
3. Find the vanishing point by marking half-way along the horizon line
4. Divide the stick man into three equal lengths to determine the braccia
5. Use a ruler to mark in the braccia intervals along the base of the fresco (A3 sheet)
6. Draw in the orthogonals from each marked braccia on the base of the fresco (A3 sheet) to the vanishing point
7. Establish the viewing distance

8. Draw in another stick man (Stick Man II) the same height as Stick Man I to establish the ideal viewing point
9. Draw a diagonal line from the top of Stick Man II's head through the vertical line which frames the left-hand side of the fresco to the first braccia
10. Where this line intercepts the vertical line which frames the left-hand side of the fresco, mark the point with an X
11. Measure the distance from the bottom left-hand corner of the fresco to the point marked X and record it
12. Go to the bottom right-hand corner of the fresco and from this point measure up the right-hand side of the exact distance obtained in Step xi and mark it with an X;
13. Draw in the first transversal by drawing a horizontal line between the two points marked X.

Three participants failed to meet the requirement to experience at least three of the four learning opportunities and were withdrawn from the treatment.

*Treatment 2.* In Treatment 2 participants received three lessons to learn the procedure which Michelangelo employed for producing the fresco painting on the Sistine Chapel ceiling. There were approximately 24 hours between each pair of learning opportunities. Prior to the start of the lessons in Treatment 2 the teacher taught the prior knowledge that participants had to acquire with respect to each component of the procedure. Participants had to demonstrate an understanding of the following concepts: *arriccio* (first rough coat of plaster); *intonaco* (final smooth coat of plaster) ; *buon fresco* (fresco painting on wet plaster) ; *secco fresco* (fresco painting on dry plaster); *giornata* (the amount of work which could be painted in one day); *pouncing* (technique for transferring the artist's drawing onto the intonaco by pricking holes in the outline for the drawing through which charcoal powder is dusted); *spolvero* (the Italian term for pouncing); cartoons (the artist's final full-scale drawing on paper); *verdaccio* (the blocking in of the bottom layers of colour); and the *Golden Hour* (the optimum period of time for applying colour to the intonaco). Only after participants had demonstrated that they understood what each concept meant were they able to sit the pre-test.

Task 2, which provided the non-pilot study content, required participants to complete a set of exercises on the scientific developments that took place in Italy during the fifteenth century.

The steps for the 13-step fresco procedure were as follows:

1. The artist makes a contract with the patron after the patron has told the artists what must be painted
2. The artist undertakes detailed drawings
3. The artist divides the work into *giornata*, the work that will be done each day
4. The artist creates the cartoons
5. The scaffolding is built
6. The *ariccio* is laid
7. The *intonaco* is laid for the *giornata*
8. The *spolvero* is completed for the *giornata*
9. Paints are prepared for the *giornata*
10. *Verdaccio* or the blocking in of the bottom layers of colour for the *giornata*
11. The modelling of form with darker and darker colours
12. The final details and strongest colours in the Golden Hour
13. The errors are corrected using *fresco a secco*.

Two participants failed to meet the requirement to experience the three learning opportunities and were withdrawn from the treatment.

*Treatment 3.* Originally Treatment 3 was planned as a two learning opportunity treatment. When Treatment 2 produced unexpected results it was decided that Treatment 3 should involve a replication of Treatment 2.

Treatment 3 involved teaching the procedure used by Piero della Francesca to produce his tempera painting of *The Baptism*. Given the unexpected results of Treatment 2 it was decided to repeat Treatment 2 with activities which were more likely to focus the participants' attention on the correct order of the steps for the procedure for creating a tempera painting.

In Treatment 3 participants received three learning opportunities with each learning opportunity separated by an inter-study interval of a single day. As in the previous two treatments, prior to the start of the pilot study lessons, the teacher taught the essential knowledge that participants had to acquire to be able to recall the procedure for producing a tempera painting. This meant that participants had to demonstrate an understanding of the following terms: *panel* (a wooden panel was used instead of a canvas for the painting surface); *tempera* (where egg yolk is used as the binder for the colour pigment); *cross-hatching* (application of the paint in crisscross strokes); cartoon (full-scale drawing on paper); pouncing (technique for transferring the artist's drawing onto the intonaco by pricking holes in the outline for the drawing through which charcoal powder is dusted); spolvero (the Italian term for pouncing); *stippling* (covering an area of the painting with dots); *under-painting* (application of the ground layers of paint); and *gesso* (plaster). After demonstrating that they understood what each term meant the participants sat the pre-test.

The teacher altered the activities so that each activity required each participant to write down the sequence of steps for producing Piero della Francesca's *The Baptism* in the correct order in each of the three learning activities. The teacher also instructed participants at the start of the treatment to put themselves in the shoes of Piero della Francesca and to think carefully about the order of the steps that they would need to take to produce his great tempera painting.

For Task 2, the task which provided the non-pilot study content, participants were required to complete a set of questions on the historical context, style and iconography of Leonardo da Vinci's *Mona Lisa*.

The steps of the procedure for producing Piero della Francesca's *The Baptism* were as follows:

1. Glue together two wooden panels which have a vertical grain



2. Glue a cloth over the panels to give the plaster something to lock onto
3. Cover the cloth with up to 8 layers of gesso
4. Cover the panel with charcoal and scrape this off until all traces of black charcoal dust have disappeared to produce a suitable flat surface
5. Complete the preliminary plan
6. Undertake the mathematical calculations needed to produce the underlying geometry of the composition and the perfectly proportioned figures
7. Create the full-scale cartoons
8. Transfer the images on the cartoons to the panel's surface using spolvero (pouncing)
9. Prepare the paints by mixing the coloured powder with egg and water
10. Block in the base colours such as the base layer of green for the figures
11. Model the forms using cross-hatching
12. Apply the fine details using techniques such as stippling which is in evidence in Jesus' beard
13. Glue gold leaf onto the angels' wings and over the part of the sky immediately above Jesus

Four participants failed to meet the requirement to experience the three learning opportunities and were withdrawn from the treatment.

## RESULTS

### **Inter-marker Agreement**

Accuracy checks were conducted by a retired teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of agreements and disagreements, and then multiplying the result by 100. For Pilot Study 1 inter-marker agreement was 100 per cent.

### **Retention of Previously Unknown Procedures**

In this pilot study each of the procedures in the three treatments had 13 steps. Participants needed to complete at least 11 of the 13 steps in the correct in order to achieve the 80 per cent correct criterion used by Nuthall. Post-test results are shown in Figure 5.

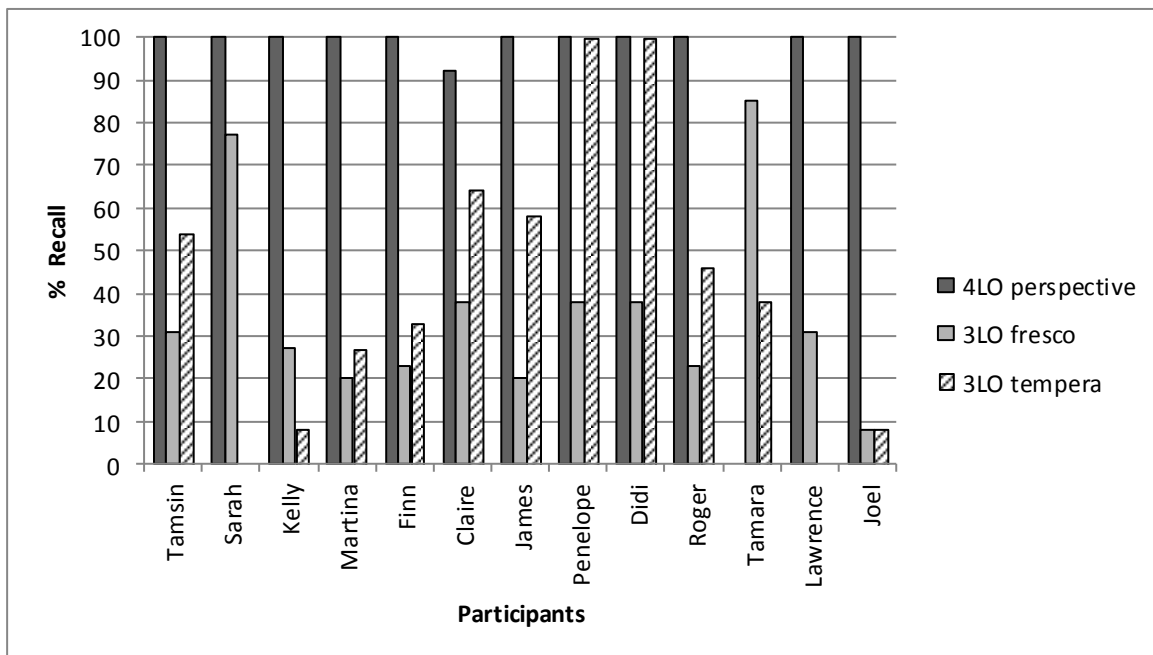


Figure 5. Percentage of steps performed correctly and in the correct order by each participant in each of the three treatments (perspective grid, fresco, and tempera) in Pilot Study 1.

*Treatment 1.* Three participants did not experience at least three of the four learning opportunities and were withdrawn from the treatment. None of the 12 participants who sat the pre-test were able to produce any of the steps required to create a single-point perspective grid for a fresco. Of the 12 participants who had received the four learning opportunities and sat the post-test all but one (Claire) performed every single step of the thirteen step procedure correctly. In step 12 of the procedure Claire did not measure the distance from the bottom right-hand corner to the mark of the first transversal adequately and, as a consequence was judged to have performed the penultimate step of the procedure incorrectly.

*Treatment 2.* Two participants had to be withdrawn from the treatment because they did not experience the three learning opportunities. In the pre-test 10 of the 13 participants failed to produce even the first step of the fresco procedure. Only three participants were able to produce any steps; Kelly produced the first two steps, while Martina and James produced the first three steps. The participants fared little better in the post-test. Only two participants were

able to recall more than half of the steps of the procedure in the correct order. Sarah produced 10 of the 13 steps and Tamara recalled 11 of the 13 steps. Eleven participants recalled five or less steps of the procedure in the correct order.

*Treatment 3.* Four participants were withdrawn from the treatment because they did not experience the three learning opportunities. While the remaining eleven participants did not experience any difficulty in learning the prerequisite knowledge for the tempera procedure, the pre-test revealed that the participants had little understanding of the procedure for producing a tempera painting. In the pre-test Claire and Martina produced the first two steps of the procedure, five others produced just the first step, and the remaining four were not able to produce a single step. On the post-test only two participants, Penelope and Didi, were able to produce the 13 steps of the procedure in the correct order. As in the fresco post-test most participants experienced great difficulty in recalling the steps in the correct order. Kelly (8%) and Joel (8%) managed to recall just one step in the correct order while four others (Martina 27%, Finn 33%, Roger 46%, and Tamara 38%) recalled less than half the steps in the correct order.

## DISCUSSION

The pilot study ...is not a necessary first step in experimental design. It is an unplanned consequence of experimentation that is carried out without sufficient knowledge of the important variables. Pilot experiments occur as frequently as they do because there are so many poorly understood factors involved, singly and in combination, in most behavioural research (Sidman, 1960, p.233).

Treatment 1 provided a replication of the results of Treatment 1 in Experiment 3 with both different subjects and different content. The fact that all 14 participants scored over 80 per cent in their post-tests provides supporting evidence of the predictive validity of the Nuthall model for the teaching of a new procedure when participants receive three or four learning opportunities.

However, the results of Treatment 2 in this Pilot Study 1 failed to replicate the results of Treatment 2 in Experiment 3. The participants' results revealed that while most of the participants could recall the majority of the steps, they experienced great difficulty in recalling the steps in the correct order. The most obvious explanation for the poor performance appears to be the different nature of the two procedures. Whereas the procedure in Treatment 1 required participants to produce something tangible as a result of performing each step of the procedure, the procedures in Treatments 2 and 3 required participants only to describe a series of steps in order. The procedure used in Treatment 1 was sufficiently well designed to enable each step in the procedure to act as a strong cue for the next step. In addition, once participants had completed the procedure in Treatment 1 they had tangible evidence of a final product. Having a very clear understanding of what a procedure will produce is likely to aid participants in correcting any errors they might make when trying to recall the procedure. While the Treatment 2 post-tests revealed that most of the participants could recall the steps, there was no real understanding of the importance of the order of the steps. As a consequence several participants failed to note that if they placed the final layer of the plaster (intonaco) over the detailed drawings on the wall then they would no longer be able to see the detailed drawings and would be left with nothing but a blank white wall. Because the participants in Treatment 3 did not have to produce a final product but only a list of actions in order, the same failure occurred on the Treatment 3 post-test. As a result of having a clear idea of what it was that they could create with the procedure in Treatment 1, the participants who carried out the procedure in Treatment 1 are likely to have been much more aware of their errors and thus better able to make self-corrections. Furthermore, in Treatment 1 it is more likely that each step in the procedure acted as a cue for the next step for participants than in the other two treatments.

The results of Treatments 2 and 3 do not constitute a disconfirmation of the Nuthall model with respect to the number of learning opportunities needed to learn a new procedure. The results do, however, indicate that learning to perform a new procedure and learning to describe a new procedure are two different kinds of learning outcomes. Although unexpected, the results of Pilot Study 1 suggest that a larger number of learning opportunities may be required in order to learn to describe a sequence of known steps in the correct order than are required to learn to perform a sequence of known steps in the correct order. The results also signal that Nuthall's rule may not apply when the learning outcome involves learning to describe a complex 13-step procedure.

## CHAPTER 7

### EXPERIMENT 4

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES BEYOND NUTHALL'S CRITICAL TWO DAY TIME WINDOW IN SETS OF YEAR 13 ART HISTORY LESSONS

Nuthall's analysis of his initial three descriptive studies led him to conclude that not only the number of learning opportunities but also the inter-study interval between those learning opportunities were crucial in determining whether or not a student learned and remembered new factual knowledge (Nuthall, 1993). Nuthall found that, when a student experienced an inter-study interval of more than two days between any pair of the four learning opportunities, then that student learned and remembered less. In his three subsequent studies the two-day inter-study interval between learning opportunities continued to play a critical role in determining what a student would and would not learn and remember in self-contained Science and Social Studies units. Nuthall argued that the two day inter-study interval functioned as a *time window* (Nuthall, 1999a). While Nuthall's theory with respect to the time window has important implications for classroom teaching practice it has never been tested experimentally.

#### AIM

The aim of Experiment 4 was to measure the effects on the retention of to-be-learned knowledge items when the inter-study intervals between pairs of learning opportunities are greater than Nuthall's two-day time window.

#### METHOD

##### **Participants and Setting**

Participants for the two treatments in Experiment 4 were the same participants who took part in Experiment 2, i.e. a mixed ability and mixed gender class of 23 Year 13 (Grade

12) Art History students from a Christchurch secondary school. Informed consent was obtained from participants and their parents using the procedures outlined in Chapter 2. Teaching and testing was conducted in the participants' normal classroom with their normal Art History teacher. The classroom contained a data projector.

### **Lesson Content**

As in Experiments 1 and 2, each experimental lesson required participants to complete two tasks with Task 1 always containing the experimental content and Task 2 always containing the non-experimental content. Again participants did not know that the lessons contained both experimental and non-experimental content. Both the experimental and non-experimental content came directly from the Year 13 Art History curriculum.

As in Experiment 2, the experimental participants studied the context, iconography and style of selected 15<sup>th</sup> Century Italian Renaissance paintings. Each treatment focused on two paintings with one providing the experimental content and the other the non-experimental content.

### **Testing Procedures**

Pretesting and retention testing were undertaken using the procedures described in Chapter 2.

### **Inter-study Intervals**

The aim of Experiment 4 was to generate a set of treatments in which at least one inter-study interval (ISI) violated Nuthall's 2-day time window. The actual arrangement of ISIs was dependent upon the school timetable for the Art History periods for this class.

### **Teaching Procedures and Learning Activities**

The four learning activities for this experiment were exactly the same as those used in Experiment 2. In order these were the Comprehension Activity, the Question Generation Activity, the Annotation Activity, and the Jumbled Letter Activity. The order of the activities

remained unchanged across the two treatments. As in the previous experiments, the teacher-generated text used for the Comprehension Activity contained each of the test items expressed as a proposition. This text was issued to each participant for every experimental lesson to ensure that, at all times, participants had access to all the information they needed to complete the activities.

*Feedback Procedures.* Feedback was provided at the end of each learning activity using the oral feedback procedure described in Chapter 2.

*Controlling the number of learning opportunities.* The procedures employed to control the number of learning opportunities were those described in Chapter 2.

## **Experimental Treatments**

*Treatment 1.* Treatment 1 consisted of four experimental lessons on the context, iconography and style for the painting *The Last Judgement* by Michelangelo. Participants received four learning opportunities using a 2-day/5-day/2-day schedule of inter-study intervals (ISI). According to this schedule there was an inter-study interval of two days between the first and second learning opportunities, an inter-study interval of five days between the second and third learning opportunities, and an inter-study interval of two days between the third and fourth learning opportunities. The first and third inter-study intervals conformed to the Nuthall model while the second inter-study interval violated the model by three days. The subject matter for the non-experimental content was the context, style and iconography for *The Disputa* by Raphael.

*Treatment 2.* Treatment 2 consisted of four experimental lessons on the context, iconography and style for the painting *Supper at Emmaus* by the Mannerist artist Jacopo Pontormo. Participants received four learning opportunities using a 3-day/4-day/3-day schedule of inter-study intervals (ISI). According to this schedule there was an inter-study interval of three days between the first and second learning opportunities, an inter-study



interval of four days between the second and third learning opportunities, and an inter-study interval of three days between the third and fourth learning opportunities. In this treatment all three inter-study intervals between the four learning opportunities violated the Nuthall two-day time window. The subject matter for the non-experimental content was the context, style and iconography for *The Resurrection* by Piero della Francesca.

Unfortunately an outbreak of illness was chiefly responsible for the fact that 12 of the 23 participants had to be withdrawn from the experiment because they did not receive at least three of the four learning opportunities in either Treatment 1 or Treatment 2.

## RESULTS

### **Inter-marker Agreement**

Accuracy checks were conducted by a fellow teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total of agreements and disagreements, and then multiplying the result by 100. For Experiment 4 inter-marker agreement was 100 per cent.

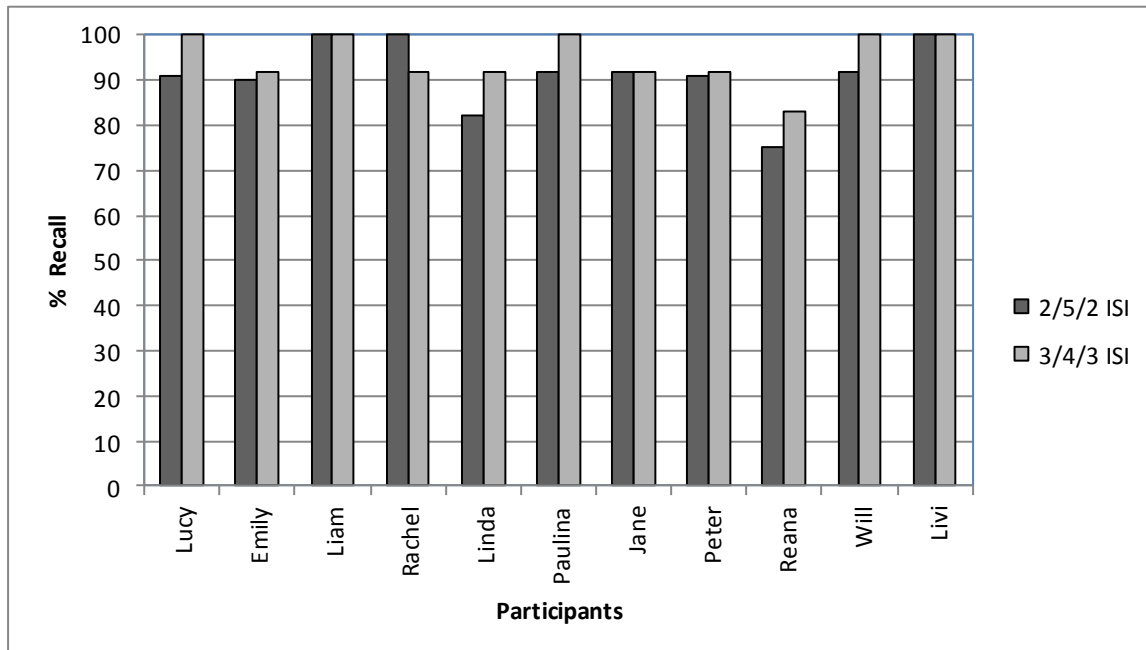
### **Retention of Previously Unknown Knowledge Items**

For the results of participants to be included in this analysis, participants had to have received at least three of the four learning opportunities in the two experimental treatments in Experiment 4. Questions which a participant answered correctly on their pre-test were not scored on that participant's post-test. The scores of each of the participants on unknown knowledge items in each of the two treatments are shown in Figure 6.

*Treatment 1.* The five day inter-study interval between the second and third learning opportunities in Treatment 1 had very little effect on the ability of the participants to learn and remember the previously unknown knowledge items about Michelangelo's *Last Judgement*. Of the 11 participants who received three or four learning opportunities in

Treatment 1 only one, Reana (75%), failed to score at least 80 per cent. Scores ranged from 75 per cent to 100 per cent.

### Retention of Previously Unknown Knowledge Items



*Figure 6.* Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 2-day/5day/2-day and 3-day/4-day/3-day schedules of inter-study intervals

*Treatment 2.* In Treatment 2 all three inter-study intervals between the learning opportunities violated Nuthall's two-day time window. For the 11 participants who received at least three learning opportunities in Treatment 2, scores ranged from 83 per cent to 100 per cent. Seven participants, Reanna (83%), Will (100%), Emily (92%), Linda (92%), Lucy (100%), Paulina (100%), and Peter (92%), scored better in Treatment 2 than they had in Treatment 1, while the other participants gained the same percentage in their Treatment 2 post-tests as they had done in their Treatment 1 post-tests.

## DISCUSSION

The fact that only 11 from a class of 23 participants experienced at least three of the four experimental lessons in each of the two experimental treatments reveals one of the true challenges of conducting educational research in the classroom. The very busy nature of the school in which the experiment was conducted meant that on occasions there were only quite small windows of opportunity for conducting experiments. Unfortunately the timing of this experiment coincided with an outbreak of illness in the school which resulted in less than half of the class being able to meet the criteria for participation in the experiment.

The results of Treatment 1 and Treatment 2 throw doubt on Nuthall's claim that an inter-study interval of more than two days between learning opportunities reduces the likelihood of students learning and remembering new knowledge items. Treatment 1 violated Nuthall's two-day time window between the second and third days by three days without interfering with the ability of most participants to learn and remember the unknown knowledge items for the painting *The Last Judgement*. Although Treatment 2 violated the two-day time window between each of the learning opportunities this failed to prevent any of the participants from scoring at least 80 per cent in their retention tests.

Overall the two treatments in Experiment 4 produced 21 disconfirmations and just a single confirmation of Nuthall's 2-day rule. It might be argued that this unexpected result may have occurred as a result of the operation of some uncontrolled or imperfectly controlled variable. This question can only be addressed by seeking to determine whether the present results can be replicated.

## CHAPTER 8

### EXPERIMENT 5

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES BEYOND NUTHALL'S CRITICAL TWO DAY TIME WINDOW IN SETS OF YEAR 12 ENGLISH LESSONS

The results of Experiment 4 throw doubt on Nuthall's claim that no more than two days can elapse between successive learning opportunities if long-term retention is to occur. In Treatment 1 of Experiment 4 a five - day inter-study interval between the second and third learning opportunities did not impair the ability of most students to learn and remember a new set of knowledge items about Michelangelo's *The Last Judgement*. When all three inter-study intervals in Treatment 2 of Experiment 4 violated the Nuthall 2-day time window by employing a 3-day/4-day/3-day schedule of inter-study intervals then this too did not prevent most students from learning and remembering a new set of knowledge items about Pontormo's *Supper at Emmaus*. Because these results were quite unexpected it was decided to see whether the results of Experiment 4 could be replicated with a different group of students studying different content.

#### AIM

The aim of Experiment 5 was to conduct a systematic replication of Experiment 4 using different participants and different subject matter.

#### METHOD

##### **Participants and Setting**

For the three treatments in Experiment 5 the participants were students in a mixed ability and mixed gender class of 21 Year 12 (Grade 11) English students at a Christchurch secondary school. The class consisted of six boys and fifteen girls with an average age of 16 years. Informed consent was obtained from participants and their parents using the

procedures described in Chapter 2. The participants experienced both the teaching and testing for all three treatments in Experiment 5 in their normal classroom with their usual English teacher. The classroom was equipped with a data projector.

### **Lesson Content**

In Experiment 5, as in the previous experiments, participants were required to complete two tasks in each 50 minute period. Task 1 contained the experimental content and Task 2 contained non-experimental content. The experimental content and the non-experimental content came from the Year 12 English curriculum. The participants were not aware which content was experimental content and which content was non-experimental content. For the experimental content in Experiment 5 participants studied a series of poems as part of a poetry unit for the short text section of the Year 12 English curriculum. Each of the three experimental treatments focused on a single poem. For the non-experimental content participants were required to complete two literature essays for their novel study and a comprehension exercise.

### **Testing Procedures**

Pretesting was undertaken using the procedures described in Chapter 2. Retention was tested using the procedures described in Chapter 2. Only those items not answered correctly on the participant's pre-test were scored on each participant's retention test.

### **Inter-study Intervals**

The aim of Experiment 5 was to generate a set of treatments in which at least one ISI violated Nuthall's 2-day time window. The actual arrangement of ISIs was dependent upon the school timetable for the English periods for this class.

### **Teaching Procedures and Learning Activities**

The learning activities for the three treatments in this experiment were the Comprehension Activity, the Question Generation Activity, the Jumbled Letter activity and a

new activity, the Interview Activity. For the Interview Activity participants had to imagine that they were the poet who had written the poem which was the subject of that treatment and, as the poet, they were required to write their responses to a set of questions about the poem.

*Feedback Procedures.* Feedback was provided at the end of each learning activity using the oral feedback procedure described in Chapter 2.

*Controlling the number of learning opportunities.* The procedures taken to control the number of learning opportunities were described in Chapter 2.

## **Experimental Treatments**

*Treatment 1.* Experiment 5 consisted of three experimental treatments. In Treatment 1 participants received four learning opportunities, one in each of the first two experimental lessons and two in the third and final experimental lesson. In the three experimental lessons participants were required to learn a set of knowledge items about the famous anti-war poem from World War 1, *Dulce et Decorum Est* by Wilfred Owen. A 4-day/1-day/0-day schedule of inter-study intervals was used to separate the four learning opportunities. In other words there was a 4-day inter-study interval between the first and second learning opportunities, a 1-day inter-study interval between the second and third learning opportunities and then an inter-study interval of just a few minutes between the third and fourth learning opportunities. This meant that Activity 4 followed immediately after the completion of Activity 3. In this treatment the first of the inter-study intervals violated Nuthall's 2-day time window. The subject matter for the non-experimental content for the first two experimental lessons was an essay on the novel *Falling* by Anne Provoost. Nine participants were withdrawn from the treatment because they did not receive at least three of the four learning opportunities.

*Treatment 2.* Treatment 2 consisted of four experimental lessons about the poem *Orca* by Robinson Jeffers. Jeffers was a misanthropist and in his poem *Orca* he contrasts the

necessary killing in nature of orcas hunting seals for food with the greed and malice that motivates humans to slaughter one another in war. The 1-day/3-day/1-day schedule of inter-study intervals used to separate the four learning opportunities in this treatment provided a three day inter-study interval between the second and third learning opportunities which violated Nuthall's two day time window. The subject matter for the non-experimental content was a second essay on the novel *Falling* by Anne Provoost. The seven participants who did not receive at least three of the four learning opportunities were withdrawn from the treatment.

*Treatment 3.* Treatment 3 consisted of four experimental lessons on another poem by Robinson Jeffers entitled *Skunks*. In *Skunks* Jeffers uses the fact that a skunk's smell, which is truly diabolical when smelt at close quarters, is strangely sweet smelling when smelt from a distance to help him to convey the idea that when viewed from afar man's most terrible actions lose their horror; it's the idea that "Distance makes clean." In this treatment a 3-day/3-day/1-day schedule of inter-study intervals separated the four learning opportunities which meant that both the first and the second study intervals violated the Nuthall two-day time window. The subject matter for the non-experimental content was a comprehension exercise which was designed to provide participants with practice answering comprehension questions for the Unfamiliar Text examination paper they would be sitting later in the year. Six participants failed to attend at least three of the four lessons and were withdrawn from the treatment.

## **RESULTS**

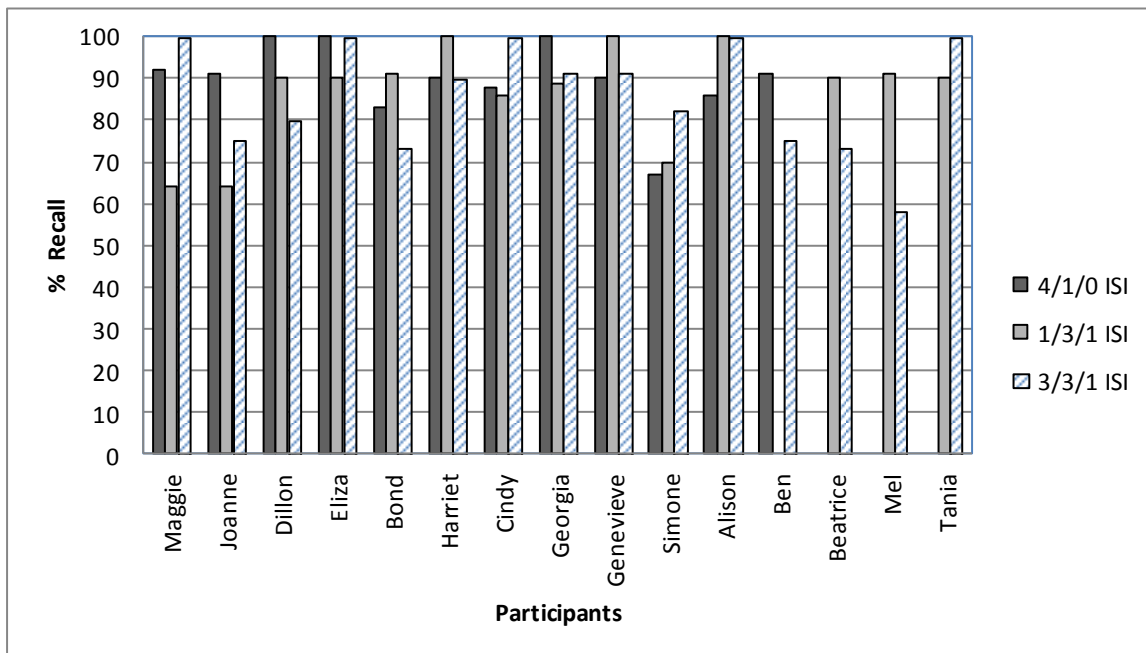
### **Inter-marker Agreement**

Accuracy checks were conducted by a teaching colleague who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of

agreements and disagreements, and then multiplying the result by 100. For Experiment 5 the inter-marker agreement was 99.9% per cent.

### Retention of Previously Unknown Knowledge Items

Only the results of participants who had received at least three of the four learning opportunities in at least two of the three experimental treatments were included in the results analysis. Once again only the questions which a participant had answered incorrectly on their pre-test were scored on that participant's post-test. Figure 7 presents the post test scores (converted to a percentage of previously unknown facts and concepts recalled in the post-test) participant by participant and treatment by treatment.



*Figure 7.* Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 4-day/1-day/0-day, 1-day/3-day/1-day, and 3-day/3-day/1-day schedules of inter-study intervals

*Treatment 1.* Twelve participants experienced three or four learning opportunities in Treatment 1. All but Simone scored more than 80 per cent in this treatment. Simone received just three of the four learning opportunities and scored 67 per cent in her post-test.



*Treatment 2.* Of the 14 participants who experienced either three or four learning opportunities in Treatment 2 only three participants, Maggie (64%), Joanne (64%) and Simone (70%), failed to score above 80 per cent.

*Treatment 3.* Fifteen participants experienced either three or four learning opportunities in Treatment 3. Five of these participants, Mel (58%), Joanne (75%), Bond (73%), Ben (75%) and Beatrice (73%), failed to score more than 80 per cent in their retention tests. Of the 10 participants who did score 80 per cent or more in their retention tests six either matched their best performance or had their best performance in Treatment 3.

## DISCUSSION

Experiment 5 systematically replicated Experiment 4 by changing both the participants and the experimental content and by varying the schedule of inter-study intervals. Treatment 1 was not a good test of Nuthall's two-day rule because even though the first inter-study interval violated the two-day time window, the remaining inter-study intervals which separated the final three learning opportunities did not. Given this schedule of inter-study intervals the Nuthall model predicted that participants were very likely to have learnt and remembered the new knowledge items about the poem because they received three further learning opportunities without there being an inter-study interval of more than two days between each one. The fact that this prediction proved correct provides yet further supporting evidence of Nuthall's model with respect to the number of learning opportunities students need to experience in order to remember new knowledge items. However, as with Treatments 1 and 2 in Experiment 4, the results of Treatment 2 and Treatment 3 in Experiment 5 throw doubt on Nuthall's claim that a two day inter-study interval constitutes a maximum time window.

Treatment 3 produced the worst overall performance across the participants. Five of the fifteen participants who experienced Treatment 3 scored less than 80 per cent in the post-test.

Perhaps part of the explanation for this is the misconception created by having two very similar questions in the pre-test and post-test. Question 6 asked “What is an aphorism?” and Question 10 asked “What is an allusion?” Only two participants failed to recall what an aphorism was but eight participants failed to recall what an allusion was. Compounding the problem was the fact that Question 11 required participants to give an example of an allusion from the poem. All the participants who scored below 80 per cent experienced problems with Questions 10 and 11.

The confusion created by having two easily confused items in a set of to-be-learned knowledge items highlights how even minor events can influence what students will and will not learn and remember in the classroom. The extent of the confusion created in the five participants who failed to score more than 80 per cent in Treatment 3 was a salient reminder that closely similar knowledge items “should not be taught together” (Church, 1999b, p. 62).

The fact remains, however, that in 21 of the 29 post-tests completed by the 15 participants in Treatments 2 and 3 of Experiment 5, participants answered 80 per cent or more of their retention tests items correctly. The results provide a second failure to confirm the Nuthall claim that this level of retention only occurs when the time elapsing between successive learning opportunities is no more than two days.

## CHAPTER 9

### EXPERIMENT 6

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO THREE TO SEVEN DAYS IN SETS OF YEAR 13 ENGLISH LESSONS

Experiments 4 and 5 measured the effects on retention of experimental treatments which deliberately extended the inter-study interval beyond the two-day inter-study interval which Nuthall claimed to be a time window. These longer inter-study intervals failed to impair the ability of most of the participants in these experiments to learn and remember new sets of knowledge items. The results of both Experiment 4 and Experiment 5 throw considerable doubt on the validity of the two-day time window component of the Nuthall model of the learning and remembering process. The results of Experiments 4 and 5 raise the question, “Just how far can the inter-study intervals between pairs of learning opportunities be extended before most of students fail to learn and remember at least 80 per cent of a set of new knowledge items?” The present experiment investigates this question by systematically replicating Experiment 5 using older participants, longer inter-study intervals, and lessons in Shakespearean drama.

#### AIM

The first aim of Experiment 6 was to try to replicate the results of Experiments 4 and 5 using different lesson content, different participants and longer inter-study intervals. The second aim was to try to establish the length of the time window which is necessary for the learning and recall of sets of knowledge items for secondary students (that is in high school).

## METHOD

### **Participants and Setting**

Participants for Experiment 6 were a mixed gender class of twenty Year 13 (Grade 12) English students at a Christchurch secondary school. There were ten boys and ten girls with an average age of 17 years. Unlike in previous experiments this class was not a truly mixed ability class. On the basis of their results from the previous year's National Certificate of Educational Achievement (NCEA) Level 2 English exam, the top performing English students in Year 13 at this school were selected at the start of their Year 13 year to join a Scholarship English class leaving the other English students to be assigned to one of three non-streamed Year 13 English classes. Included in the participants' Year 13 class were two students, David and Taylor, who had been in an alternative English programme in Year 12 (Grade 11) because their English was below the standard required for a normal Year 12 class. Both students had to get a special dispensation to allow them to enter the normal Year 13 English programme. Several participants in this group, including Taylor, had had their lives seriously disrupted by the earthquakes. While the experimental lessons were taking place after-shocks continued to be felt. The informed consent of these students was obtained using the Informed Consent procedures described in Chapter 2.

Damage caused by the Christchurch earthquakes to the participants' normal classroom meant that the teaching and testing took place in an alternative classroom in the school. Participants were instructed by their usual classroom teacher throughout the course of the experiment.

### **Lesson Content**

In each 50 minute experimental lesson participants again completed two tasks. Task 1 provided the experimental content and Task 2 provided the non-experimental content. As in previous experiments, participants were not aware which content was experimental and

which content was non-experimental as they did not know that they were participating in an experiment. The experimental and non-experimental content both came directly from the Year 13 English curriculum. For the experimental content participants studied *Othello* in preparation for their NCEA paper on Shakespeare and for the non-experimental content participants completed a series of literature essay paragraphs on the play *Death and the Maiden* in preparation for their NCEA paper on Non-Shakespearean Drama.

### **Testing Procedures**

The pretesting and post-testing testing procedures described in Chapter 2 and used in preceding experiments were used in Experiment 5.

### **Inter-study Intervals**

The aim of Experiment 6 was to generate a set of treatments in which all of the inter-study intervals in each treatment violated Nuthall's 2-day time window. As with the previous two experiments, the actual arrangement of ISIs was dependent upon the school timetable for the English periods for this class.

### **Teaching Procedures and Learning Activities**

The four learning activities for Experiment 6 were exactly the same as those used in Experiment 4. These were the Comprehension Activity, the Question Generation Activity, the Interview Activity and the Jumbled Letter Activity. For the Interview Activity in Treatment 1 participants had to imagine that they were Shakespeare himself in order to answer a series of interview questions about *Othello*. For the Interview Activity in Treatment 2 participants had to put themselves in the shoes of the arch villain, Iago, to answer his interview questions.

The procedures taken to control the number of learning opportunities experienced by each participant and the procedures used to provide feedback on participants' responses were the same as those described in Chapter 2.

## **Experimental Treatments**

*Treatment 1.* Experiment 6 consisted of two experimental treatments. Treatment 1 consisted of four experimental lessons on knowledge items about the first half of the play *Othello*. Participants received four learning opportunities using a 3-day/3-day/3-day schedule of inter-study intervals. All three inter-study intervals violated Nuthall's two day time window. The subject matter for the non-experimental content for the four experimental lessons was a literature essay on the drama *Death and the Maiden* which the participants had been studying.

*Treatment 2.* In Treatment 2 participants received four experimental lessons on knowledge items about the second half of the play *Othello*. Participants experienced four learning opportunities using a 4-day/6-day/7-day schedule so that, once again, all three inter-study intervals violated the Nuthall time window of two days. The subject matter for the non-experimental content was a literature essay for the class's film study on *Schindler's List*.

Six of the 20 participants failed to meet the requirement to experience at least three of the four learning opportunities in both Treatment 1 and Treatment 2 and were withdrawn from the experiment.

## **RESULTS**

### **Inter-marker Agreement**

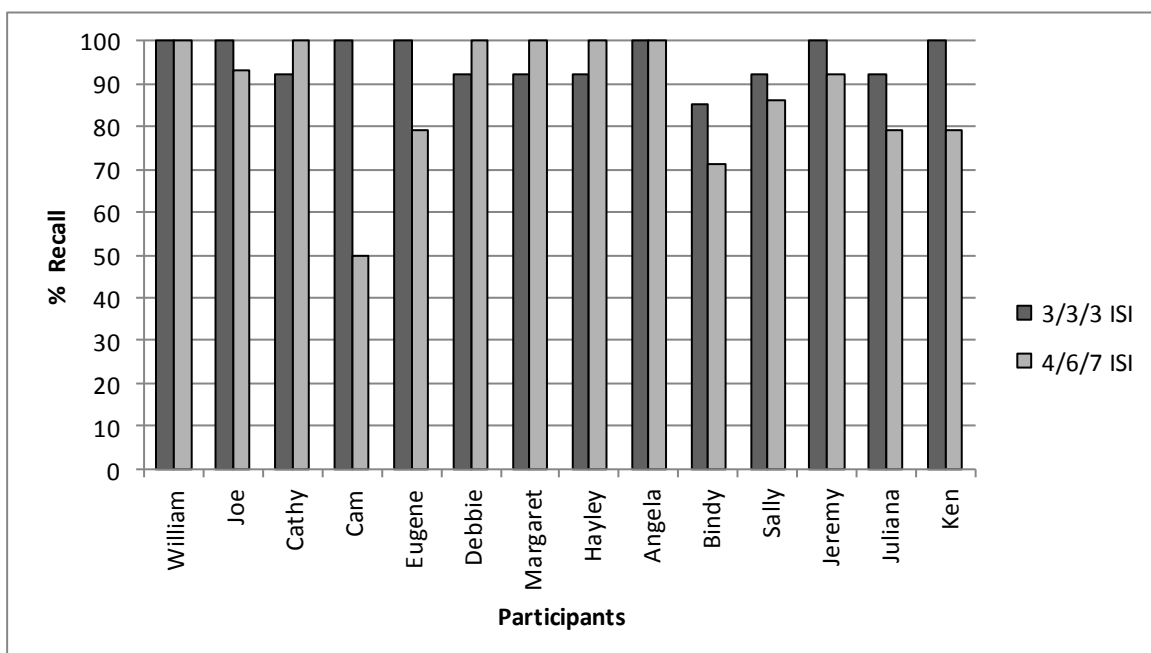
Accuracy checks were conducted by a fellow teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of agreements and disagreements, and then multiplying the result by 100. For Experiment 6 inter-marker agreement was 99.9% per cent.

## Retention of Previously Unknown Knowledge Items

As in previous experiments, questions which a participant answered correctly on their pre-test were not scored on that participant's post-test. Post test scores (converted to a percentage of previously unknown knowledge items recalled in the post-test) are presented in Figure 8, participant by participant and treatment by treatment.

*Treatment 1.* In Treatment 1 where fourteen participants experienced four learning opportunities using a 3-day/3-day/3-day schedule of inter-study intervals the lowest post-test score was Bindy's with 85 per cent. Seven participants scored 100 per cent.

*Treatment 2.* In Treatment 2 where the same fourteen participants were presented with four learning opportunities using a 4-day/6-day/7-day schedule of inter-study intervals 9 of the 14 participants scored above 80 per cent and 6 of the 9 scored 100 per cent. Of the five participants who scored less than 80 per cent three scored 79%. Only Bindy (71%) and



*Figure 8.* Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 3-day/3-day/3-day and 4-day/6-day/7-day schedules of inter-study intervals

Cam (50%) failed to get close to 80 per cent. William and Angela scored 100 per cent in both treatments. While the percentage recall of eight of the participants declined in Treatment 2, four participants, Margaret (100%), Hayley (100%), Debbie (100%) and Cathy (100%) actually performed better in Treatment 2 with its significantly longer inter-study intervals than they had in Treatment 1.

## DISCUSSION

Experiment 6 sought to replicate Experiment 5 using longer inter-study intervals with different participants and different lesson content. The 3-day/3-day/3-day schedule of inter-study intervals used in Treatment 1 produced the highest levels of retention achieved in any of the experiments to date. In Treatment 2 the 4-day/6-day/7-day schedule of inter-study intervals appears to have had a mixed impact on the ability of the participants to learn and remember a set of previously unknown knowledge items about the play *Othello* with four of the participants performing better in Treatment 2 than they had in Treatment 1 while eight performed less well in Treatment 2 than in Treatment 1. However, most of the participants continued to score either very close to or above 80 per cent. Only Cam's performance appears to have been significantly affected by the increased lengths of the inter-study intervals. Having gained a score of 100 per cent in Treatment 1, Cam managed to score just 50 per cent in Treatment 2. The reasons for this are not clear. Among the many possible reasons for this is the possibility that the poor retention was due to the fact that Cam received just three of the four learning opportunities. Cam missed the second learning opportunity, and hence experienced an inter-study interval of 10 days between his first and second learning opportunity. Ken was the other participant who received only three of the four learning opportunities. Ken had scored 100 per cent in Treatment 1 but only 79 per cent in Treatment 2.



Overall the strength of the performances of this class came as somewhat of a surprise. In particular the performances of both Joe and Eugene were unexpected. They were among the weakest Year 13 students the classroom teacher had taught and yet both performed well in the two experimental treatments. Nuthall (1999a) provides a possible answer as to why this group of individuals performed so strongly in both treatments. Nuthall found that a narrative can provide an especially memorable learning opportunity. In Nuthall's Study 6 (Nuthall, 1999a; 2000a) where students were studying a Science unit on Antarctica, Nuthall discovered, through interviews with his five student subjects, that talks given by two guest speakers about their personal experiences in Antarctica had proven to be memorable experiences for his students. Nuthall argued that the five student subjects in this study used their memories of the narratives given by the two speakers to construct their responses to specific questions in their 12 month post-test. The fact that the teacher-generated texts for both Treatments 1 and 2 in Experiment 6 were written like a narrative may help to explain the surprisingly strong performances from the participants in this Year 13 class.

Nevertheless, it is clear from the results of Experiment 6 that inter-study intervals of 6 and 7 days did not have an obviously adverse effect on the learning and remembering of these 14 participants – once again failing to provide support for Nuthall's claim that retention requires a time window of no more than two days between successive opportunities to study previously unknown facts.

Overall the two treatments in Experiment 6 produced 23 disconfirmations and just five confirmations of Nuthall's two-day rule.

## CHAPTER 10

### EXPERIMENT 7

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO THREE TO SEVEN DAYS IN SETS OF YEAR 12 ENGLISH LESSONS

The results of Experiment 6 provided a further failure to find support for Nuthall and Alton-Lee's claim that retention depends on inter-study intervals of no more than 2 days. The high retention levels of the participants in Treatment 2 of Experiment 6 which employed a 4-day/6-day/7-day schedule of inter-study intervals suggest that if a time window for the recall of previously unknown knowledge items does exist then it may be as long as seven days. A new teaching year provided the opportunity to test further the generalisability of the results of Experiment 6 by replicating that experiment with a new class of Year 12 (Grade 11) English students, different content and even longer inter-study intervals.

#### AIM

The first aim of the present experiment was to try to replicate the results of Experiment 6 using different students and different content. The second aim was to measure the effects on retention of using a schedule of inter-study intervals in which all three inter-study intervals were at least six days long.

#### METHOD

##### **Participants and Setting**

Participants for Experiment 7 were a mixed ability and mixed gender class of Year 12 English students at a Christchurch secondary school. The class contained eight boys and eight girls with an average age of 16 years. For the most part, the participants were a group of conscientious and mature students who had the ability to work independently for extended periods of time. Like many of their peers in Christchurch schools, this group had developed a

good deal of resilience in the face of the massive disruptions resulting from the Christchurch earthquakes and this enabled them to continue to function effectively in the classroom despite experiencing frequent aftershocks. The informed consent of the participants and their parents for this experiment was obtained using the procedures described in Chapter 2. While teaching and testing was conducted by the participants' usual English teacher, it did not take place in the participants' normal classroom due to damage caused by the earthquakes. Both treatments in this experiment took place in an alternative classroom within the school. The classroom was equipped with a television and a video player. Participants sat in small groups of three to six students as they had in their previous classroom.

### **Lesson Content**

The experimental and non-experimental content was drawn from the Year 12 English curriculum. Participants again completed two tasks during the 50 minute experimental lessons with Task 1 always providing the experimental content and Task 2 providing the non-experimental content. Participants had no knowledge of which content in each experimental lesson was experimental content and which content was not. For the experimental content participants studied two poems as part of their preparation for the Level 2 short text National Certificate of Educational Achievement (NCEA) examination. Each experimental treatment focused on a single poem. For the non-experimental content participants worked on literature essays for the study they had just completed on the set novel, *Falling* by Anne Provoost.

### **Testing Procedures**

The pretesting procedures and the post-test procedures were as described in Chapter 2. Only those test items which a participant was unable to answer in the pre-test were scored on the participant's post-test.

## **Inter-study Intervals**

The aim of Experiment 7 was to generate a set of treatments in which, once again, all of the inter-study intervals in each treatment violated Nuthall's 2-day time window. Like the previous experiments, the actual arrangement of ISIs was dependent upon the school timetable for the English periods for this class.

## **Teaching Procedures and Learning Activities**

The four learning activities in Experiment 7 were the same as those used in Experiments 4 and 5. These activities were the Comprehension Activity, the Question Generation Activity, the Interview Activity, and the Jumbled Letter Activity. Feedback on the participants' responses was provided by using the oral feedback procedure described in Chapter 2 and the procedures used to control the number of learning opportunities were the same as those described in Chapter 2.

## **Experimental Treatments**

*Treatment 1.* Experiment 7 consisted of two experimental treatments. Treatment 1 consisted of four experimental lessons about the poem *Skunks* by the misanthropic poet, Robinson Jeffers. It will be recalled that in his poem *Skunks* Robinson Jeffers shows how with the passage of time the deeds of history's greatest tyrants become sanitized so that "Distance makes clean." Participants received four learning opportunities using a 4-day/3-day/4-day schedule of inter-study intervals. This meant that all three inter-study intervals violated the *less than* two day time window which Nuthall claimed to be essential for retention. The non-experimental content consisted of a literature essay on the novel the participants had been studying.

*Treatment 2.* In Treatment 2 participants received four experimental lessons to learn a set of to-be-learned knowledge items about the poem *Labyrinth* by Elliot Richman. In this anti-war poem, which is a clever allusion to Wilfred Owen's great anti-war poem from World

War One, *Dulce et Decorum Est*, Richman describes what it is like going down into a tunnel in search of the enemy from the perspective of an American soldier during the Vietnam War. Participants received four learning opportunities using a 6-day/7-day/6-day schedule of inter-study intervals. With this schedule all three inter-study intervals again violated Nuthall's two-day time window. Task 2 (the non-experimental content) required participants to complete a second essay on their novel.

Three of the 16 participants failed to meet the requirement to experience at least three of the four learning opportunities in both Treatment 1 and Treatment 2 and were withdrawn from the experiment.

## RESULTS

### **Inter-marker Agreement**

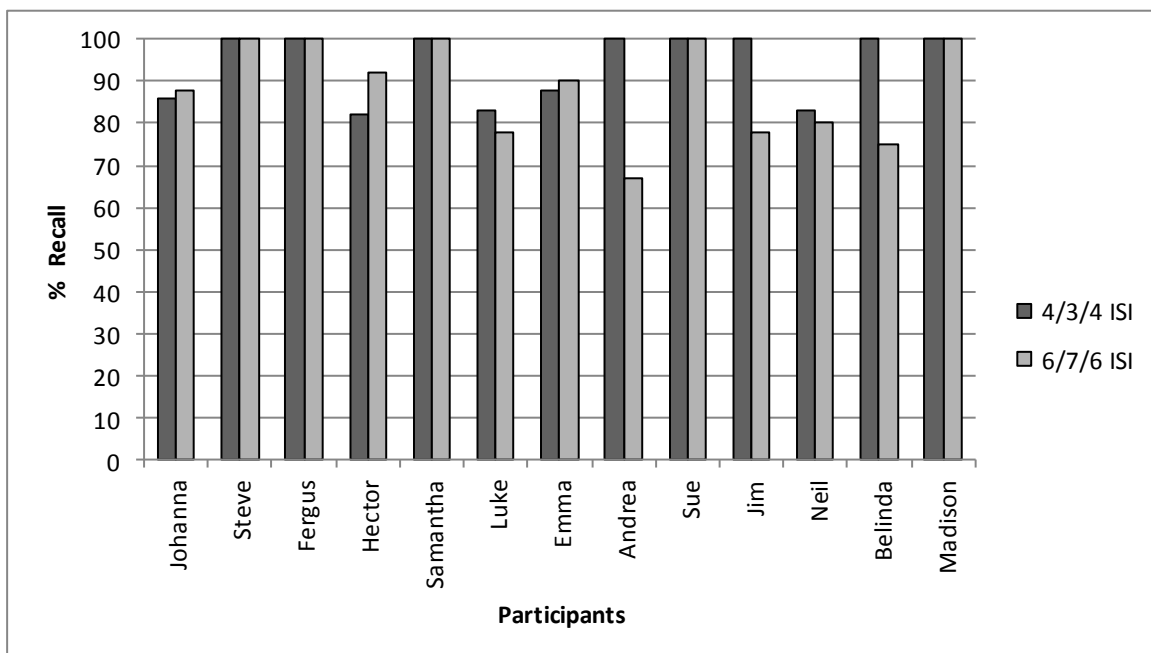
Accuracy checks were conducted by a colleague who check-marked the scoring of all responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of agreements and disagreements, and then multiplying the result by 100. For Experiment 7 inter-marker agreement was 100% per cent.

### **Retention of Previously Unknown Knowledge Items**

Questions which a participant answered correctly on their pre-test were not scored on that participant's post-test. Post test scores (converted to a percentage of previously unknown knowledge items recalled in the post-test) are presented in Figure 9, participant by participant and treatment by treatment.

*Treatment 1.* In Treatment 1 participants experienced four learning opportunities using a 4-day/3-day/4-day schedule of inter-study intervals. All of the participants scored at least 80 per cent in the post-test. Eight of the thirteen participants scored 100 per cent in the post-test.

*Treatment 2.* In Treatment 2 13 participants were presented with four learning opportunities using a 6-day/7-day/6-day schedule of inter-study intervals. Four participants, Luke (78%), Andrea (67%), Belinda (75%), and Jim (78%) failed to score at least 80 per cent. Five participants scored less in Treatment 2 than they had in Treatment 1 while two participants, Emma (90%) and Hector (92%), actually achieved a higher score in Treatment 2 than in Treatment 1.



*Figure 9.* Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 4-day/3-day/4-day and 6-day/7-day/6-day schedules of inter-study intervals

## DISCUSSION

Experiment 7 sought to replicate Experiment 6 by changing the participants, the content and the length of the inter-study intervals. Treatment 1 presented participants with the second most challenging schedule of inter-study intervals used in any of the experiments to date. However, every participant who received at least three of the four learning opportunities managed to score 82 per cent or more in the post-test.

Treatment 2 presented participants with the most challenging schedule of inter-study intervals employed in any of the experiments. In Treatment 2 there was almost a week between each of the four learning opportunities. Despite the considerable increase in the length of the inter-study intervals from Treatment 1 to Treatment 2 the performances of 10 of the 13 participants changed very little. Only the performances of Belinda (-25%), Andrea (-33%), and Jim (-22%) declined significantly as a result of increasing the inter-study intervals in Treatment 2.

Four participants failed to score at least 80 per cent in their post-tests in Treatment 2. Jim and Luke answered seven of the nine questions they had not known in the pre-test correctly in the post-test and so each scored 78 per cent in their post-tests. Belinda made three errors and Andrea four errors and thus scored 75 per cent and 67 per cent respectively in their retention tests. Question 1 which asked for the name of the poet and Question 11 which asked why the poet had made a reference to mustard gas in his poem were the hardest questions judging by the number of participants who were not able to answer these questions correctly.

Interestingly three of the four participants who failed to score at least 80 per cent made errors with both questions and the fourth answered Question 11 incorrectly. As a result of an oversight in the construction of the final activity participants were denied a fourth learning opportunity to learn why the poet had made a reference to mustard gas in his poem. The fact that Belinda missed the final experimental lesson meant that she actually only received two learning opportunities to learn the answer to Question 11 and this may explain why she answered Question 11 incorrectly.

Overall the two treatments in Experiment 7 produced 22 disconfirmations and just four confirmations of the Nuthall two-day rule. The results of Experiment 7 provide further empirical evidence that Nuthall's two day inter-study interval does not constitute a time window when it comes to recalling new factual knowledge with secondary school aged

participants. The ability of most of the participants in both treatments to learn and remember sets of new knowledge items with two challenging schedules of inter-study intervals provides further evidence that the time window, if it exists, might be as great, or greater, than a week.



## CHAPTER 11

### EXPERIMENT 8

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO MORE THAN A WEEK IN SETS OF YEAR 13 ART HISTORY LESSONS

In Experiment 7 inter-study intervals of up to seven days failed to prevent Year 12 participants from learning and remembering sets of factual information about two 20<sup>th</sup> Century poems. The results of Experiment 7 throw further doubt on Nuthall and Alton-Lee's claim that retention is unlikely to occur when more than two days elapse between successive learning opportunities. The new academic year brought a new set of Year 13 Art History students and with them came the opportunity to test the generalizability of the results of Experiment 7 by replicating that experiment with different subjects, different content and further increases in inter-study interval length.

#### AIM

The first aim of Experiment 8 was to try to replicate the results of Experiment 7 using different students and different content. The second aim was to measure the effects on retention of increasing the inter-study interval to 8 or 9 days.

#### METHOD

##### **Participants and Setting**

Participants for the present experiment were a class of 20 mixed ability and mixed gender Year 13 (Grade 12) Art History students. The class contained ten boys and ten girls with an average age of 17 years. The Christchurch earthquakes had had a devastating effect on several of the participants in this class and aftershocks continued to make their presence felt during some of the experimental lessons. Informed consent was obtained from participants and their parents using the procedures described in Chapter 2. Damage caused by

the Christchurch earthquakes meant that both the teaching and testing in this experiment took place in a more earthquake resilient classroom in the school and not in the participants' normal classroom. The participants were taught by their usual Art History teacher during the experimental lessons.

### **Lesson Content**

In each 50 minute experimental lesson the experimental content provided by Task 1 and the non-experimental content provided by Task 2 came directly from the Year 13 Art History curriculum. For Task 1 participants studied two paintings, one by Piero della Francesca in Treatment 1 and one by Sandro Botticelli in Treatment 2. For Task 2 participants studied two Leonardo da Vinci portraits in Treatment 1 and Raphael's fresco paintings in the Papal apartments in Treatment 2. The participants had no idea they were participating in an experiment and, as a consequence, could not have differentiated between the experimental and the non-experimental content.

### **Testing Procedures**

Pretesting and retention testing for each treatment were undertaken using the procedures described in Chapter 2.

### **Inter-study Intervals**

The aim of Experiment 8 was to generate a set of treatments in which, once again, all of the inter-study intervals in each treatment violated Nuthall's 2-day time window. Like the previous experiments, the actual arrangement of ISIs was dependent upon the school timetable for the Art History periods for this class.

### **Teaching Procedures and Learning Activities**

Each activity in Experiment 8 provided each participant with four three-term learning interactions with every unknown knowledge item. The four activities in this experiment were

the same as those used in Experiments 2 and 3. These were the Comprehension Activity, the Question Generation Activity, the Annotation Activity, and the Jumbled Letter Activity.

The feedback procedures which were used were those described in the Chapter 2 and the procedures taken to control the number of learning opportunities each participant received were the same as those described in Chapter 2.

### **Experimental Treatments**

*Treatment 1.* Experiment 8 consisted of two experimental treatments. Treatment 1 consisted of four experimental lessons on a set of knowledge items about Piero della Francesca's painting *The Misericordia*. Participants received four learning opportunities using a 5-day/6-day/4-day schedule of inter-study intervals. The actual intervals were dictated in part by the school timetable for Year 13 Art History. Again, all three inter-study intervals violated Nuthall's two day time window. For the non-experimental content in each of the experimental lessons participants studied the context, meaning and style of two of Leonardo da Vinci's portraits, *Lady with an Ermine* and *Ginevra de Benci*.

*Treatment 2.* Treatment 2 consisted of four experimental lessons on a set of knowledge items about Sandro Botticelli's *Birth of Venus*. Participants received four learning opportunities using an 8-day/9-day/5-day schedule of inter-study intervals. For the non-experimental content participants studied the context of Raphael's frescoes in the Papal apartments of the Vatican.

Four of the 20 participants failed to meet the requirement to experience at least three of the four learning opportunities in both Treatment 1 and Treatment 2 and were withdrawn from the experiment.

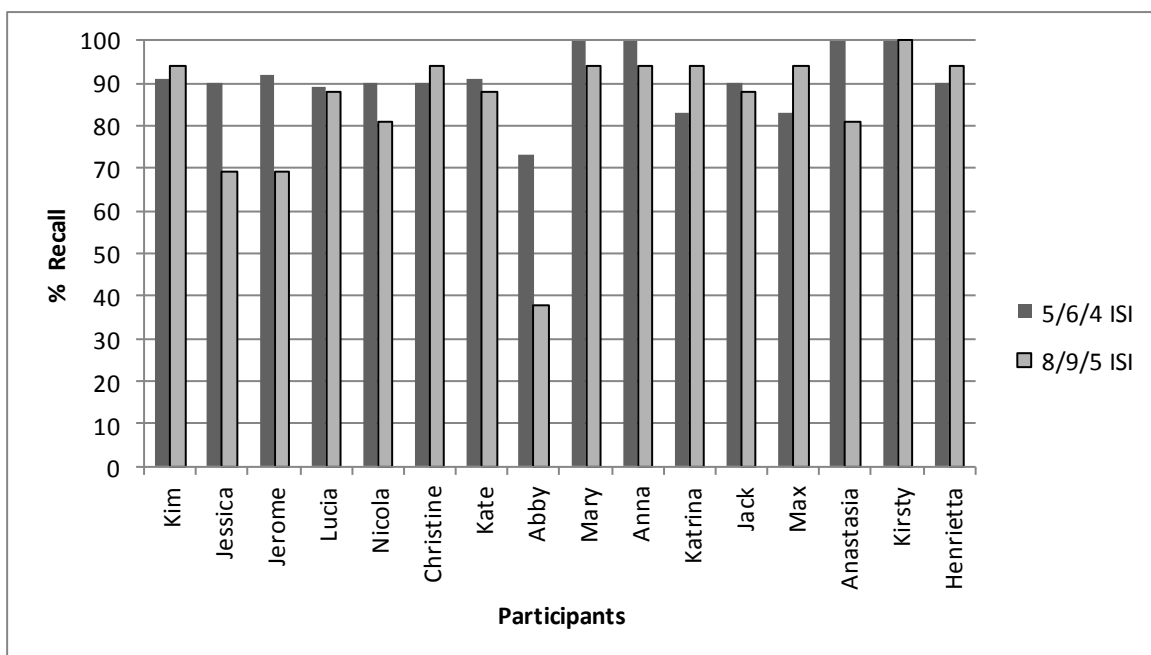
## RESULTS

### Inter-marker Agreement

Accuracy checks were conducted by a fellow teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of agreements and disagreements, and then multiplying the result by 100. For Experiment 8 inter-marker agreement was 99.9% per cent.

### Retention of Previously Unknown Knowledge Items

Questions which a participant answered correctly on their pre-test were not scored on that participant's post-test. The post-test scores for each participant in each of the two treatments are presented in Figure 10.



*Figure 10.* Percentage of unknown knowledge items correctly recalled by each of the participants following the completion of treatments with 5-day/6-day/4-day and 8-day/9-day/5-day schedules of inter-study intervals

*Treatment 1.* Treatment 1 consisted of four experimental lessons about the iconography, style and context of *The Misericordia*. When the sixteen participants experienced four learning opportunities using a 5-day/6-day/4-day schedule of inter-study intervals only one participant, Abby (73%), failed to achieve a score of at least 80 per cent in the post-test. Mary, Anastasia, Kirsty and Anna all scored 100 per cent in the post-test. Katrina (83%), Abby (73%), Anna (100%) and Christine (90%) experienced only three of the four learning opportunities. Abby, Anna and Christine all missed the second learning opportunity which meant that there was an inter-study interval of 11 days between their first and second learning opportunities. Katrina missed the third learning opportunity which meant that there was an inter-study interval of 10 days between her second and third learning opportunities.

*Treatment 2.* Treatment 2 consisted of four experimental lessons on knowledge of the style, context and iconography of Botticelli's *Birth of Venus*. Participants experienced four learning opportunities using an 8-day/9-day/5 day schedule of inter-study intervals. Three participants, Abby (38%), Jerome (69%) and Jessica (69%) failed to score at least 80 per cent in their retention tests. It will be recalled that Abby was the only participant not to have scored at least 80 per cent in Treatment 1. Her score of just 38 per cent was the lowest percentage recall score of any participant in any of the experimental treatments in which participants received three or four learning opportunities. As in Treatment 1, Abby experienced just three of the four learning opportunities. She missed the final learning opportunity which meant that she sat her retention test 11 days after her final learning opportunity. Anastasia (81%), Kirsty (100%), Anna (94%) and Lucia (88%) also experienced just three of the four learning opportunities. Anastasia, Kirsty and Lucia each missed the second learning opportunity which meant that there was an inter-study interval of 17 days between their first and second learning opportunities. Anna missed the third learning

opportunity and this meant that there was an inter-study interval of 14 days between her second and third learning opportunities.

## DISCUSSION

Despite the lengthy inter-study intervals in both treatments, 13 of the 16 participants achieved scores of 80 per cent or more on their two retention tests. It was interesting to note the performances of those participants who experienced just three of the four learning opportunities in either Treatment 1 or Treatment 2. With the notable exception of Abby every participant who experienced just three of the four learning opportunities still managed to score 80 per cent or more in their retention tests despite the very long inter-study intervals of 10 and 11 days between pairs of learning opportunities in Treatment 1 and inter-study intervals of 14 and 17 days between pairs of learning opportunities in Treatment 2.

Abby's very poor performance was not unexpected. Abby's motivation at the time when the experiment was undertaken was at a low ebb. She had fallen behind in her work in several of her subjects, failed to complete several internal assessments and had been under pressure from teachers to improve both her attendance at classes and her performance in those classes. In addition Abby's engagement in class activities generally was quite variable. Abby was often off task in class and it was not unusual for Abby to seek clarification about a teacher's instructions after her peers had settled to a task that the teacher had just spent some time explaining.

The fact that inter-study intervals of eight and nine days did not prevent most of the participants from learning and remembering a set of previously unknown knowledge items about Botticelli's *Birth of Venus* suggests that if a time window does exist it is probably greater than 9 days for this age group.

Overall the two treatments in Experiment 8 produced 28 disconfirmations and just four confirmations of the Nuthall two-day rule.

## CHAPTER 12

### EXPERIMENT 9

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES TO MORE THAN TWO WEEKS IN SETS OF YEAR 13 ENGLISH LESSONS

Experiment 8 failed to find an outer limit to the time window beyond which most students would fail to learn and remember a new set of knowledge items despite the fact that inter-study intervals of 8 and 9 days had been used. In Experiments 4 to 8 relatively small increments in the inter-study intervals between successive learning opportunities had been made. In all five experiments, however, students had succeeded in learning and remembering most of the knowledge items studied. While these experiments indicated fairly clearly that learning and remembering does not depend on the spacing of learning opportunities with “no more than a 2-day gap” (Nuthall, 1999a, p. 305), as claimed by Nuthall and Alton-Lee, they failed to find any kind of outer limit to the time window between learning opportunities. This led to the decision to increase the inter-study intervals to the point where it might be expected that most students would fail to learn and remember new knowledge items. It was hoped that such an experiment might identify a possible end point from which future researchers, who were interested in determining the true outer limit of the hypothesized time window, could work back from.

#### AIM

The first aim of Experiment 9 was to try to replicate the results of Experiment 5 using exactly the same experimental content, but different participants and much longer inter-study intervals. The second aim was to try to identify an inter-study interval which is sufficiently long to prevent a majority of students from retaining a set of previously unknown knowledge items.

## METHOD

### **Participants and Setting**

Participants for this experiment were a mixed gender class of 16 Year 13 English students at a Christchurch secondary school. There were 12 boys and four girls with an average age of 17 years. As with the Year 13 (Grade 12) English class used in Experiment 5, this class was not a truly mixed ability class. On the basis of their results from the previous year's National Certificate of Educational Achievement (NCEA) Level 2 English examination, the top performing English students in Year 13 at this school were selected at the start of their Year 13 year to join a Scholarship English class leaving the other English students to be assigned to one of three non-streamed Year 13 English classes. Compared to the Year 13 English class who took part in Experiment 6, the class for Experiment 9 was a much less capable class. Whereas the class of participants in Experiment 6 had two of their number who had just managed to gain entry into the normal Year 13 English course the present class had half a dozen students who, while they had gained entry to the normal Year 13 English course, were more suited to the Year 13 Alternative English class because their written language skills were so poor. Two of these participants had been diagnosed with specific learning difficulties. During the period of the present experiment participants had to contend with further aftershocks and a severe snow storm. Informed consent was gained from the participants and their parents using the Informed Consent procedures described in Chapter 2. Teaching and testing was conducted in the participants' normal classroom with their normal English teacher. During the lessons described below the class was seated in groups containing between three and six students. The classroom was equipped with a data projector.

### **Lesson Content**

Participants were required to complete two tasks in each 50 minute experimental lesson. Task 1 provided the experimental content and Task 2 provided the non-experimental



content. Participants were not able to differentiate between the two types of content because they were oblivious to the fact that they were participating in an experiment. The experimental and non-experimental content both came from the Year 13 English curriculum. For the experimental content participants studied *Othello* so as to provide an introduction to the teaching of the *Othello* unit proper which was scheduled to begin later in the academic year. For the non-experimental content participants completed literature essays and research questions on the play *Death and the Maiden* and the film *Schindler's List* in preparation for their external examinations and a research internal assessment.

### **Testing Procedures**

The pretesting and the retention test procedures used in previous experiments were also used in Experiment 9. These procedures were described in Chapter 2.

### **Inter-study Intervals**

The aim of Experiment 9 was to generate a set of treatments in which, once again, all of the inter-study intervals in each treatment violated Nuthall's 2-day time window. The actual arrangement of ISIs was dependent upon the school timetable for the English periods for this class.

### **Teaching Procedures and Learning Activities**

The four learning activities for Experiment 9 were the Comprehension Activity, the Question Generation Activity, the Interview Activity and the Jumbled Letter Activity. As in Experiment 5 the Interview Activity in Treatment 1 required participants to climb into the shoes of Shakespeare himself in order to answer a series of interview questions about *Othello* from the playwright's perspective. Again, as in Experiment 5, the Interview Activity in Treatment 2 required participants to imagine themselves as being the arch villain, Iago, in order to answer interview questions from his perspective.

The feedback procedures described in Chapter 2 continued to be used. The procedures

taken to control the number of learning opportunities were the same as those described in Chapter 2.

### **Experimental Treatments**

*Treatment 1.* Experiment 9 consisted of two experimental treatments with Treatment 1 consisting of four experimental lessons on knowledge items about the first half of the play *Othello*. Participants received four learning opportunities using a 7-day/7-day/7-day schedule of inter-study intervals. For Task 2, where participants worked with non-experimental content, participants worked on a series of literature essays and research questions on the drama *Death and the Maiden* by the Chilean playwright, Ariel Dorfman.

*Treatment 2.* In Treatment 2 participants received four experimental lessons on knowledge items about the second half of the play *Othello*. The participants experienced four learning opportunities using a 14-day/28-day/20-day schedule of inter-study intervals. A number of factors conspired together to prevent the operation of a schedule of learning opportunities with more consistent inter-study intervals. The most significant limiting factor was the need to complete the treatment before the scheduled teaching of *Othello* in Term 3. This left just 11 school weeks, the two final weeks of Term 1 and the nine weeks of Term 2, in which to complete Treatment 2. Complicating the situation was the three week holiday break which separated Term 1 and Term 2. For Task 2 participants worked on a series of literature essays and research questions on Steven Spielberg's epic film *Schindler's List*.

Four of the 16 participants failed to meet the requirement to experience at least three of the four learning opportunities in both Treatment 1 and Treatment 2 and were withdrawn from the experiment.

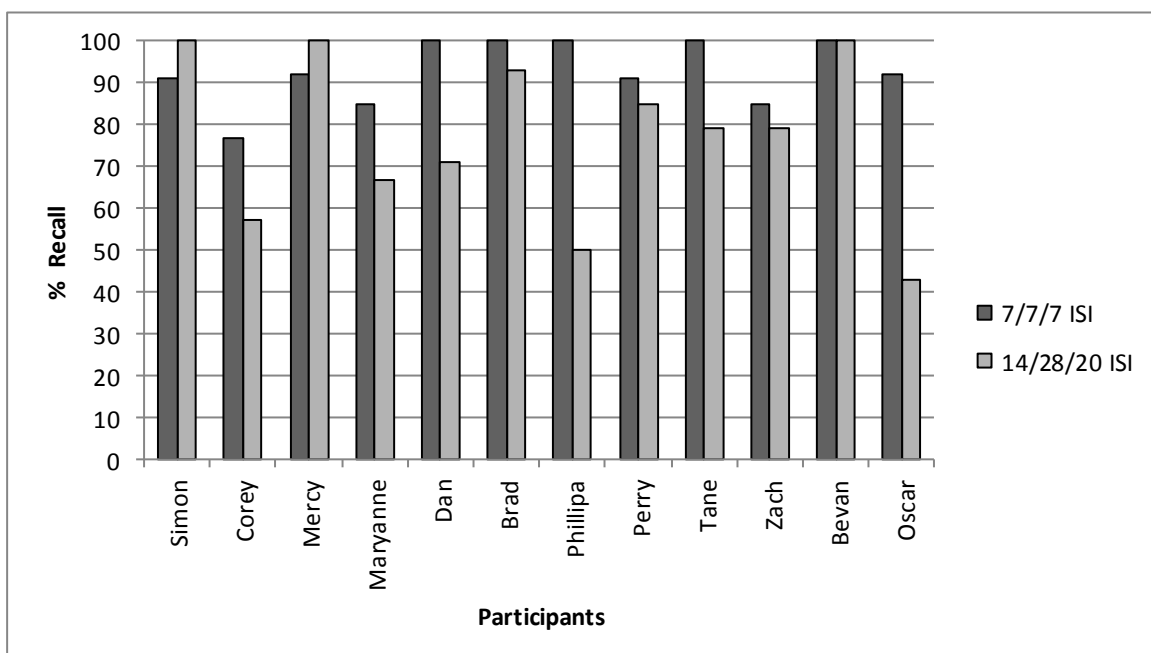
## RESULTS

### Inter-marker Agreement

Accuracy checks were conducted by a fellow teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of agreements and disagreements, and then multiplying the result by 100. For Experiment 9 inter-marker agreement was 99.9% per cent.

### Retention of Previously Unknown Knowledge Items

As in previous experiments questions which a participant answered correctly on their pre-test were not scored on that participant's retention test. Retention test scores (converted to a percentage of previously unknown knowledge items recalled in the post-test) are presented in Figure 11, participant by participant and treatment by treatment.



*Figure 11.* Percentage of unknown knowledge items recalled correctly by each of the participants following the completion of treatments with 7-day/7-day/7-day and 14-day/28-day/20-day schedules of inter-study intervals.

*Treatment 1.* In Treatment 1 participants experienced four learning opportunities using a 7-day/7-day/7 day schedule of inter-study intervals. Of the 12 participants who sat the pre-test all but Corey (77%) scored over 85 per cent. Three-quarters of the participants scored 90 per cent or more in their post-tests.

*Treatment 2.* Five of the 12 participants achieved scores of more than 80 per cent in their retention tests. Three participants, Simon, Mercy and Bevan, scored 100 per cent. Bevan had also scored 100 per cent in the Treatment 1. Simon's result was especially surprising as he was one of the weakest English students in the class and a student who really struggled with motivation in English. The other two participants to gain more than 80 per cent were Brad (93%) and Perry (85%). Both Tane and Zach scored 79 per cent in their post-tests. Tane's performance was surprising given that he had missed the second learning opportunity as a result of being away on a Physical Education trip. This meant there was actually an inter-study interval of 42 days between his first and second learning opportunity. Two other participants who had been on that same Physical Education trip and experienced a 42 day inter-study interval between their first and second learning opportunities did not fare as well. Maryanne with 67 per cent and Oscar with 43 per cent were among the seven participants who failed to score at least 80 per cent. The other five participants who failed to score at least 80 per cent were Zach (79%), Tane (79%), Dan (71%), Corey (57%), and Phillipa (50%).

## DISCUSSION

Experiment 9 replicated Experiment 6 using the same experimental content but different participants and a more widely spaced schedule of inter-study intervals. The 7-day/7-day/7-day schedule of inter-study intervals used in Treatment 1 of Experiment 9 produced very similar results to those achieved in Treatment 1 (3-day/3-day/3-day schedule of ISI) and Treatment 2 (4-day/6-day/5-day schedule of ISI) of Experiment 6. The consistently high retention rates of the participants in Treatment 1 were somewhat surprising

given the fact that a number of the class found English very challenging. The very mixed results from Treatment 2 which used a 14-day/28-day/20-day schedule of inter-study intervals do no more than hint at the possibility that a 28-day inter-study interval might prove to be a reasonable starting point from which a future researcher might wish to work back from in an attempt to identify the length of any time window which might exist for Level 3 NCEA students when learning a new set of facts and definitions. However, the results of Treatment 2, in particular the extremely high retention levels of Simon, Bevan, Mercy and Brad, also suggest the possibility that the construct of a time window during which long term memories are formed may be an empty construct – at least when it comes to learning during the teenage years.

When the number of disconfirmations and confirmations of Nuthall's 2-day rule in Experiment 9 is added to the number of disconfirmations and confirmations from the previous five experiments it gives a cumulative total of 131 disconfirmations and 30 confirmations of Nuthall's 2-day rule. It can be argued, however, that the results of Treatment 2 of Experiment 9 should not be included in the cumulative total given that the very challenging schedule of inter-study intervals was deliberately chosen to try to prevent students from recalling the knowledge items. If the results of Treatment 2 are not included this gives a cumulative total of 126 disconfirmations and just 23 confirmations. This means that 84.6 per cent of the results from 12 separate experimental treatments failed to provide confirmation of the Nuthall claim that repeated exposures to new knowledge content must occur within a 2-day time window if that content is to be both learned and remembered.

## CHAPTER 13

### EXPERIMENT 10

#### EFFECTS ON RETENTION OF EXTENDING THE INTER-STUDY INTERVAL BETWEEN LEARNING OPPORTUNITIES BEYOND TWO DAY DAYS DURING THE TEACHING OF A NEW MATHEMATICS PROCEDURE WITH YEAR 10 STUDENTS

Experiment 3 and Pilot Study 1 investigated whether the first part of the Nuthall model for learning and remembering (which specifies the number of learning opportunities which are required) could be generalized to the learning of unknown procedures. Treatments 1 and 2 in Experiment 3 and Treatment 1 in Pilot Study 1 provided strong support for the validity of the first component of the Nuthall model. Treatments 3 and 4 in Experiment 3, however, did not. It will be recalled that in both treatments most of the participants were able to learn and remember the previously unknown procedures for drawing a line graph and a hyperbola after receiving just two learning opportunities. In Experiment 10 the generalizability of the second component of the Nuthall model (which states that retention depends on there being no more than a two day inter-study interval between any pair of learning opportunities) was examined with respect to the learning of new procedures.

#### AIM

The first aim of the experiment was to try to replicate with younger students the findings from Experiment 3 where participants received four learning opportunities to learn and remember a new Mathematics procedure under tightly controlled conditions. The second aim of the experiment was to measure the effects on the retention of a new Mathematics procedure when the inter-study intervals between pairs of learning opportunities are extended beyond Nuthall's critical two-day time window.

## METHOD

### **Participants and Setting**

The participants for the three experimental conditions in Experiment 10 were a group of 13 students from a Year 10 (Grade 9) Mathematics class. This group consisted of two girls and eleven boys who were of average to above average ability in Mathematics. None of the participants were aware that they were taking part in an experiment. The number of participants who entered each treatment and the number of participants who completed each treatment are given in the Results section below. Informed consent was obtained from participants and their parents using the procedures described in Chapter 2. Teaching and testing were conducted in the participants' usual Mathematics classroom with their usual classroom teacher.

### **Lesson Content**

Participants completed several tasks in each 50 minute experimental lesson. Task 1 consisted either of a warm up activity aimed at developing participants' fluency in basic multiplication facts or a short test on Algebra. Task 2 always required participants to work with the experimental content. The experimental content for Task 2 was exactly the same experimental content as that used in the first three treatments of Experiment 3 and it came from a self-contained unit on drawing graphs. Participants were required to learn three 12 to 14 step procedures for graphing a parabola, a circle and a line. Each of the three experimental treatments focused on one of these graphs. Task 3 always required participants to complete a set of exercises on a different unit of work.

### **Testing Procedures**

*Pre-tests.* Sitting pre-tests at the start of new units in Mathematics was a common occurrence for these Year 10 students. Participants sat separate pre-tests for each of the three treatments one to five days prior to the start of the experimental lessons for that treatment.

The pre-tests used for this experiment were exactly the same as those used in the first three treatments of Experiment 3.

*Retention tests.* Participants sat a retention test which was exactly the same as the pre-test a week after the final learning opportunity. Participants who were not present at the post-test could still be included in the experiment provided they sat the post-test within an 8 to 15 day period following the final experimental lesson. Participants who did not sit the post-test within this 8 to 15 day period were removed from the experiment.

### **Inter-study Intervals**

The aim of Experiment 10 was to generate a set of treatments in which at least one of the inter-study intervals in each treatment violated Nuthall's 2-day time window. The actual arrangement of ISIs was dependent upon the school timetable for the Mathematics periods for this class.

### **Teaching Procedures and Learning Activities**

*Pre-teaching of the prerequisite component responses.* To identify the key component responses and the precise sequence of steps which go up to make each procedure, the teacher completed a task analysis of each of the three graphing procedures. The teacher then created a schedule for teaching the key component knowledge and skills. To try to prevent participants from intuitively working out the correct sequence of steps required to complete the procedure selected for a particular treatment, the teacher taught the component knowledge and skills in a random order and spread out over several lessons. Only those participants who demonstrated that they had learned and remembered the key component skills in a short test were able to sit the pre-test. One participant in Treatment 1 and two participants in Treatment 3 failed to meet this requirement and were withdrawn from these treatments.

*Procedures for the experimental lessons.* Each set of four learning activities with the experimental content in each treatment in Experiment 10 required participants to perform the



procedure which was the subject of that experimental treatment. At the start of each experimental lesson each participant was issued with an A3 sheet which contained both the instructions for drawing the graph selected for that treatment and graph paper for drawing the graph. Detailed instructions were supplied for the first two experimental lessons in each experimental treatment. The instructions issued for the third and fourth learning opportunities in Treatment 1 and for the third learning opportunity in Treatment 2 were much less detailed. This was done so that participants in Treatments 1 and 2 would be forced to try to recall some of the steps which had been deliberately missed out in the instructions for these learning opportunities. Participants who could not recall all the key steps of a procedure using abbreviated instructions were issued with the instructions that had been provided for the first experimental lesson. In other words all participants had the opportunity to refer to the full set of instructions should they need to use them during the third and fourth learning opportunities.

The teacher began the first experimental lesson in each treatment by explaining to the participants that they were going to learn a procedure for drawing a particular graph. After identifying the graph the teacher then explained that the procedure for drawing the graph had been broken down into a series of steps on their A3 sheet and that they would need to follow the steps carefully in order to produce the graph on the graph paper supplied on the A3 sheet. The teacher determined the pace at which the graph was completed in the first experimental lesson by requiring participants to complete each step of the procedure with him. The teacher monitored the progress of each participant carefully to make sure that each participant was able to successfully complete the procedure. For each subsequent experimental lesson the participants were able to complete the procedure at their own pace. When a participant encountered difficulties the teacher limited his assistance to either pointing to an instruction the participant had failed to follow or by asking the participant a question which was

designed to prompt the participant to find their own solution to his or her difficulty. Feedback was provided in accordance with the instructions which are described in Chapter 2 and the number of learning opportunities was controlled using the written materials procedure described in Chapter 2.

### **Experimental Treatments**

*Treatment 1.* Experiment 10 consisted of three experimental treatments. Treatment 1 was exactly the same as Treatment 1 in Experiment 3 where participants were provided with four experimental lessons to try to learn and remember how to draw a parabola graph. Participants received four opportunities to complete the 13 step procedure for drawing a parabola graph, one on each day for four successive days. Task 3 which provided the non-experimental content required participants to complete a set of exercises for an Algebra unit.

The procedure for drawing a parabola graph was broken down into its eleven key components. These were being able to: (a) recognize the formula for a parabola; (b) recognize the shape of a parabola graph; (c) locate specified co-ordinates on a graph; (d) identify the  $x$  and  $y$  axes on a graph; (e) find the  $x$  intercepts from the equation for a parabola; (f) find the mid-point between the two  $x$  intercepts ; (g) find the mirror line between two  $x$  intercepts; (g) find the  $y$  intercept from the equation for a parabola; (h) add integers; (h) multiply integers; and, finally, (i) explain the concept of symmetry.

Participants had to demonstrate in a short test that they were able to perform each of the component skills identified above before they were allowed to sit the pre-test.

It will be recalled that the 14 step procedure for drawing the parabola  $(x + 4)(x - 3)$  is as follows:

Recognise it is a parabola graph

1. Find the first  $x$  intercept by making the bracket  $(x + 4) = 0$ , i.e.  $x = -4$
2. Mark this on the  $x$  axis

3. Find the second  $x$  intercept by making the bracket  $(x - 3) = 0$ , i.e.  $x = 3$
4. Mark this on the  $x$  axis
5. Find half way between the two intercepts by first counting the squares between the two intercepts, i.e. 7 squares
6. Divide this number by 2; i.e.  $7/2 = 3.5$
7. Travel 3.5 squares from any one of the  $x$  intercepts towards the other and mark it with an X to identify the mid-point
8. Write in the co-ordinates for the mid-point next to X, i.e.  $(-0.5, 0)$
9. Draw a vertical line through the mid-point to form a line of symmetry for the parabola
10. Find the  $y$  intercept by first making  $x = 0$  in both brackets, i.e.  $y = (0 + 4)(0 - 3)$
11. Multiply  $4 \times -3 = -12$
12. Mark in the  $y$  intercept with an X
13. Draw in a symmetrical parabola which passes through both  $x$  intercepts and the  $y$  intercept.

In the final learning opportunity Jordan and Aimee asked to be able to correct their errors on a clean graph. The teacher made the decision to permit this as this was what typically would have occurred in a normal classroom activity and to deny the request and insist that the corrections be made on the sheet where the errors occurred would have produced a potentially confusing image of the parabola. In Experiment 3 it will be recalled that while completing his third learning opportunity to draw the parabola one of the participants realised he had made an error without any teacher feedback and asked for and received another piece of graph paper from the teacher so that he could complete the graph correctly. This was the only example of a student making a self-correction in Experiment 3. In this treatment Matthew and Jordan made self-corrections in the third learning opportunity and Tom made a self-correction in the fourth learning opportunity without receiving any teacher feedback. The three participants asked for and received a clean graphing sheet so that they could complete their graphs.

*Treatment 2.* In Treatment 2 participants received four learning opportunities to complete all 12 steps of the procedure for drawing a circle graph using a 4-day/3-day/4-day schedule of inter-study intervals. Prior to the beginning of the experimental lessons participants were taught the prerequisite skills and knowledge required to complete each of the key components of the procedure in a random order both at different times during a lesson and in separate lessons. Task 3, which provided the non-experimental content, required participants to complete a set of angles exercises for a Geometry unit.

The 12 step procedure for drawing a circle graph was broken down into its five key components. These were being able to: (a) recognize the formula for a circle; (b) find a point in all four quadrants of a graph using a set of co-ordinates; (c) find the square root of 9, 16, 25, 49, 64 and 81; (d) find the  $x$  and  $y$  co-ordinates from any given circle graph formula; and, (e) locate points on a graph for a given number of squares above, below, to the right or to the left of a given point on a graph.

As in Treatment 1, participants had to demonstrate that they were able to complete each of the component skills before they were permitted to sit the pre-test. Care was taken to teach the component skills and knowledge in a disjointed fashion so as to prevent the possibility of participants working out the correct sequence of steps needed to draw the circle graph for themselves. Any participants who were able to complete the procedure after being taught the component skills and knowledge were withdrawn from the experimental treatment.

The 12 step procedure for drawing the circle  $(x + 2)^2 + (y - 1)^2 = 9$  was as follows:

1. Recognise the need to draw a circle graph from the equation  $(x + 2)^2 + (y - 1)^2 = 9$
2. Find the square root of 9
3. Write down what the radius is, i.e. radius = 3
4. Find the  $x$  co-ordinate by making the bracket  $(x + 2) = 0$ , i.e.  $x = -2$
5. Find the  $y$  co-ordinate by making the bracket  $(y - 1) = 0$ , i.e.  $y = 1$
6. Locate the point  $(-2, 1)$  on the graph and mark it with an O

7. Next to the O write down the co-ordinates for the centre of the circle
8. From O go right 3 squares and mark it with an X
9. From O go up 3 squares and mark it with an X
10. From O go left 3 squares and mark it with an X
11. From O go down 3 squares and mark it with an X
12. Draw the circle graph by drawing a circle through each of the 4 points marked with an X.

*Treatment 3.* Treatment 3 consisted of four experimental lessons on how to draw a line graph. An untimely and significant snowfall, illnesses, and sports injuries to several participants interrupted the planned schedule of inter-study intervals for this treatment. Instead of the 6-day/7-day/7-day schedule of inter-study intervals which had been originally planned, an 11-day/3-day/6-day schedule of inter-study intervals was used for six participants and an 11-day/7-day/2-day schedule had to be used for the remaining two participants, Ivan and Melanie, who were eligible for the treatment. The decision was made to allow Ivan and Melanie to continue with the treatment even though they were using a different schedule of inter-study intervals because, while they experienced a longer inter-study interval between the second and third learning opportunities, the actual lengths of their 20-day schedule of inter-study intervals was similar to the 20-day schedule of the other students.

The 13 step procedure for drawing the line graph was broken down into its six key components. These were being able to: (a) recognize  $y = x$  as the formula for a line graph; (b) find the highest common factor between three numbers; (c) find the value of  $y$  in a linear equation by making  $x = 0$  and working out the answer; (d) find the value of  $x$  in a linear equation by making  $y = 0$ ; (e) change the subjects of algebraic equations; and, (f) work out the answer by making  $x$  the subject. When teaching the prerequisite skills the teacher took great care to use letters other than  $x$  and  $y$  to try to prevent the participants from associating the new teaching with graphing lines. Instead of presenting participants with the equation

$3y = 6x - 9$  which may have allowed participants to make a connection between what they were doing and the drawing of a line graph, participants received a disguised linear equation in a form such as  $3k = 6m - 9$ . The teacher took great care to teach the skills required to manipulate the algebraic terms required to find the  $x$  and  $y$  intercepts for a line graph without making any references to graphing of any sort.

Two participants who were unable to complete each of the component skills were not permitted to sit the pre-test and were withdrawn from the treatment. Three participants who were able to complete the procedure during the pretesting were also withdrawn from the treatment.

The 13 step procedure for drawing the line graph  $3y = 12x + 18$  was as follows:

1. Recognise  $3y = 12x + 18$  is a line graph
2. Divide the equation by 3 ( $y = 4x + 6$ )
3. Find  $y$  by making  $x = 0$ ;  $y = 4 \times 0 + 6$
4. Solve this;  $y = 6$
5. Write down the co-ordinates for the first point (0, 6)
6. Mark the point on the graph
7. Find  $x$  by making  $y = 0$ ;  $0 = 4x + 6$
8. Make  $x$  the subject by subtracting 6 from both sides  $-6 = 4x$
9. Divide both sides of the equation by 4 ( $-6/4 = x$ )
10. Solve this;  $x = -1.5$
11. Write the co-ordinates for the second point (-1.5, 0)
12. Mark the point on the graph
13. Draw in the line between the 2 points

## RESULTS

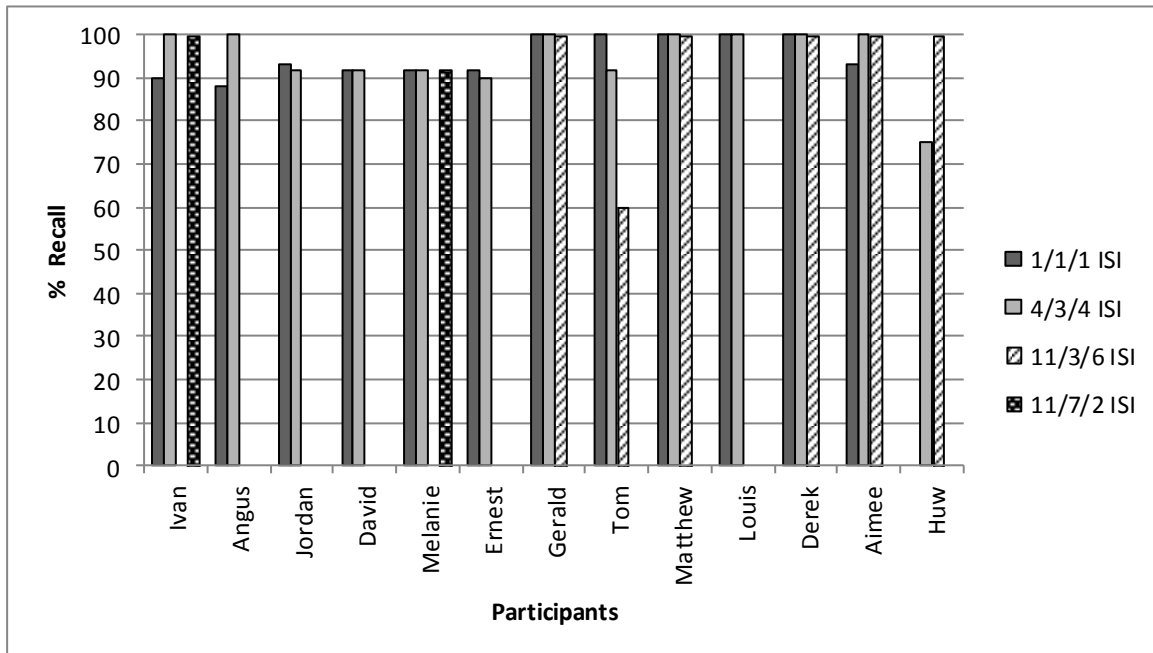
### Inter-marker Agreement

Reliability checks were conducted by a retired mathematics teacher who check-marked the scoring of all of the participants' responses in the pre-tests and post-tests. Inter-marker agreement was determined by dividing the total number of agreements by the total number of

agreements and disagreements, and then multiplying the result by 100. For Experiment 10 inter-marker agreement was 98.7% per cent.

### Retention of Previously Unknown Procedures

The results for Experiment 10 are presented in Figure 12.



*Figure 12.* Percentage of steps performed correctly and in the correct order for each participant following the completion of Treatment 1 (parabola) with 1-day/1-day/1-day inter-study intervals, Treatment 2 (circle) with 4-day/3-day/3-day inter-study intervals, and Treatment 3 (line) with either an 11-day/3-day/6-day or an 11-day/7-day/2-day inter-study intervals

*Treatment 1.* One student was withdrawn from the treatment because he did not have the prerequisite knowledge needed to sit the pre-test. None of the 12 participants who had demonstrated that they had learned the component skills for graphing a parabola and completed the pre-test were able to draw the parabola correctly in the pre-test. Angus and Derek came closest to drawing the parabola as both drew a parabola and were able to locate the  $x$  intercepts correctly. They were not, however, able to locate the  $y$  intercept. Ivan was

able to locate the two  $x$  intercepts and the mid-point between them but he could go no further than this and failed to draw a graph. Despite demonstrating that they knew that the equation  $y = x^2$  produced a parabola in a test of the component skills prior to sitting the pre-test, only six of the participants actually drew a parabola in the pre-test. Three drew a circle, one a line and two failed to draw any graph. Eleven of the 12 participants received all four learning opportunities. Ivan received only the first three learning opportunities.

In the post test all 12 participants demonstrated that they had learned the sequence of steps required to draw the graph of the parabola  $y = (x + 4)(x - 3)$ . Five participants completed all the steps in the procedure correctly. The other seven participants each made a single error. The most common error made by the participants was the failure to draw a symmetrical graph.

*Treatment 2.* All of the 13 participants in Treatment 2 qualified to sit the pre-test. In the pre-test none of the participants was able to complete the circle graph. In fact the highest score in the pre-test was just 1/12. Nine participants successfully completed every step of the procedure in the post-test. Jordan, Ernest, and Tom each missed the step which required them to write down the co-ordinates of the centre of the circle. Huw was the only participant not to perform the procedure correctly. He failed to recall the steps for finding the  $x$  and  $y$  co-ordinates and, as a result, located the wrong point for the centre of the circle. However, he was still able to produce a circle graph with the correct radius. Four participants, Aimee, Jordan, Derek and Ivan received just three of the four learning opportunities but still managed to perform the procedure accurately.

*Treatment 3.* In the pre-test three participants, Ernest, David and Angus, were able to complete the entire procedure for drawing the line graph and so they were withdrawn from the experimental treatment. Two participants, Jordan and Louis, were withdrawn from the experiment as they did not have the prerequisite knowledge and skills. As a result of



interruptions caused by snow and illness only four participants received four learning opportunities using the 11-day/3-day/6-day schedule of inter-study intervals and all four completed each step correctly in the right order in the retention test.

Two participants, Ivan and Melanie, received four learning opportunities using an 11-day/7-day/2-day schedule of inter-study intervals. Melanie made a slight division error in the second step of the procedure but thereafter completed each step of the procedure accurately. Ivan produced the line graph accurately but forgot to write down the co-ordinates of both points on the graph. Two participants, Derek and Tom, received just three learning opportunities. Although Derek missed the first learning opportunity he successfully completed each step of the procedure with a 3-day/6-day schedule of inter-study intervals. Tom, who missed the final learning opportunity due to a sports injury and thus received just three learning opportunities with an 11-day/3-day schedule of inter-study intervals, failed to complete the line graph accurately because he put the  $x$  and  $y$  values around the wrong way when writing the co-ordinates for his two points.

## DISCUSSION

Given that Treatment 1 in this experiment was a direct replication of Treatment 1 in Experiment 3 the results were expected to be very similar. This, in fact, proved to be the case. In both experiments all of the participants who received either three or four learning opportunities with 1-day inter-study intervals were able to draw the parabola and no participant made more than a single error in doing so.

Together Treatment 2 and Treatment 3 provided a total of 19 disconfirmations and just 2 confirmations of the second component of the Nuthall model which states that the inter-study interval between any pair of learning opportunities must be no more than two days. In the present experiment all but Huw were able to carry out the procedure for drawing a circle accurately using a 4-day/3-day/4-day schedule of learning opportunities in Treatment 2 and,

with the exception of Tom, all of the participants were able to successfully complete the procedure for graphing a linear equation using either an 11-day/3-day/6-day or an 11-day/7-day/2-day schedule of learning opportunities.

## CHAPTER 14

### DISCUSSION

“For generalization, psychologists must finally rely, as has been done in all older sciences, on replication” (Cohen, 1994, p. 997).

The primary aim of the experiments described in this thesis was to test the predictive validity of Nuthall’s model of the learning and remembering process. This model states that most students will learn and remember new knowledge provided they experience three to four encounters with the “complete set” of information about each item of new knowledge with no more than a two day gap between any pair of these encounters. This final chapter reviews the major findings of the present set of experiments, examines the reliability and the generality of the findings, relates the findings both to current theories of learning and memory and to previous research, identifies the main implications of the present experiments for teaching practice in schools, and suggests areas for future research.

In his descriptive studies Nuthall was not able to exercise any control over the number of encounters students had with each element of the to-be-learned content, the nature of the students’ encounters with the content or the time intervals between the students’ encounters with the content. This meant that, during the course of a series of classroom lessons, students often encountered only fragments of individual knowledge items. This led Nuthall to devise a complex procedure which enabled him to code these fragments and to count the *complete sets* of information encountered by each individual student. He was also able to measure the time intervals between each of these encounters for each student. It was this procedure that allowed him to predict with an accuracy of between 80 to 85 per cent the knowledge items which individual students would and would not learn and remember a year after the completion of the observed lessons. Nuthall’s model appears to provide teachers with an incredibly efficient and effective means of achieving durable learning. While Nuthall was

able to replicate the results of his descriptive studies six times he understood that the generality of his theory could only be determined by experimentation.

### **Overview of the Present Results**

*The number of learning opportunities required to learn and remember new knowledge items.* It will be recalled that in this research a learning opportunity was defined as a single response to a single stimulus with feedback. Using this definition of a learning opportunity the results of Experiments 1 and 2 were consistent with the predictive validity of the first component of Nuthall's model regarding the number of learning opportunities students need in order to learn and to remember new knowledge items. With few exceptions, when Year 7 and Year 13 participants received one learning opportunity each day for three or four consecutive days they scored at least 80 per cent or more in their individual post-tests of unknown knowledge items completed one week following the final learning opportunity. A significant feature of the findings was that virtually all students, regardless of their academic ability, were able to perform at the level predicted by Nuthall's model.

According to Nuthall's model, Year 7 and Year 13 participants should not be able to learn and remember a previously unknown set of knowledge items when they receive just two learning opportunities. In Experiments 1 and 2 the retention test performances of most participants declined significantly when they received just two learning opportunities. Only 4/18 participants in Experiment 1 and 3/19 participants in Experiment 2 scored at least 80 per cent in their individual post-tests in the two learning opportunities treatments. With just two learning opportunities there was a much greater variability in the post-test scores for both the Year 7 and Year 13 participants with the more academically able participants generally being least affected and the less academically able participants being most affected by the reduction in learning opportunities.

Table 1

*Number of confirmations and disconfirmations of Nuthall's 3 to 4 learning opportunity (LO) rule for the recall of unknown knowledge items for the two experiments which tested this rule*

Treatment	No of LOs	Schedule of LOs	N	No of Confirmations ≥ 80%	No of Disconfirmations < 80%
E 1 T 1 (KI) Yr 7 Sci	4	(1/1/1)	18	17	1
E 1 T 2 (KI) Yr 7 Sci	3	(1/1)	18	14	4
E 1 T 3 (KI) Yr 7 Sci	2	(1)	17	13	4
E2 T 1 (KI) Yr 13 Art Hist	4	(1/1/1)	20	20	0
E2 T 2 (KI) Yr 13 Art Hist	3	(1/1)	15	14	1
E 2 T 3 (KI) Yr 13 Art Hist	2	(1)	17	14	3

<sup>1</sup> E = Experiment, T = Treatment, KI = Knowledge Item, , Yr 7-13 = Year level at school, Sci = Science, Art Hist = Art History

Across the six experimental treatments in Experiments 1 and 2, summarised in Table 1, there were 92 confirmations and 13 disconfirmations of Nuthall's three to four learning opportunity rule for learning and remembering new knowledge items. This means that 88 per cent of the results in Experiments 1 and 2 were consistent with Nuthall's 3 to 4 learning rule for the learning and remembering of new knowledge items. It needs to be recalled that whereas Nuthall's retention tests were predominantly recognition tests, the retention tests in the present study were more demanding as they were all constructed response tests of recall.

Overall, the results of this research provide strong supporting evidence for the Nuthall model with respect to the number of learning opportunities needed to learn and remember new knowledge items.

*The number of learning opportunities needed to learn a new procedure.* The present study did not confirm Nuthall's model of the learning and remembering process with respect

to the number of learning opportunities required to learn and remember a new procedure. In Treatment 3 of Experiment 3 most of the Year 12 Mathematics participants were able to learn and remember a new procedure for graphing a linear equation after experiencing just two learning opportunities one day apart. Treatment 4 (which was a replication of Treatment 3) in which participants again received one learning opportunity on each of two consecutive days to learn the procedure for drawing a hyperbola graph appeared to provide a further disconfirmation of the Nuthall model with respect to the number of learning opportunities participants need to learn and remember a new procedure. It will be recalled that in each experimental treatment in Experiment 3 great care had been taken to teach the separate steps in the new procedure in isolation from one other. This was done to ensure that the participants knew how to perform the majority of steps in the procedure before being taught the sequence of steps needed to complete the new procedure. The results of Treatments 3 and 4 in Experiment 3 suggest that, provided a teacher takes care to ensure that his or her students know how to perform the majority of steps in a new procedure, students can learn and remember the sequence of steps in a new procedure after experiencing just two learning opportunities.

*Pilot Study 1: The learning of procedures in Year 13 Art History.* It will be recalled that the pilot study described in Chapter 6 consisted of three separate treatments. In Treatment 1 participants were required to learn and remember a procedure for creating a single perspective grid. When participants received four learning opportunities with one day separating each pair of learning opportunities their post-test performances provided strong support for Nuthall's model. However, in Treatment 2 in which participants were required to recall the procedure for painting a fresco painting, after having experienced three learning opportunities, the poor post-test performance of the participants was not predicted by the

Nuthall model. As a result of the unexpected results of Treatment 2 it was decided to repeat Treatment 2 with activities which were more likely to focus the participants' attention

Table 2

*Number of confirmations and disconfirmations of Nuthall's 3 to 4 learning opportunity (LO) rule for the recall of unknown procedures for Experiment 3 and Pilot Study 1.*

Treatment	No of LOs	Schedule of LOs	N	No of Confirmations $\geq 80\%$	No of Disconfirmations $< 80\%$
E 3 T 1 Procedure Yr 12 Maths	4	(1/1/1)	7	7	0
E 3 T 2 Procedure Yr 12 Maths	3	(1/1)	9	9	0
E 3 T 3 Procedure Yr 12 Maths	2	(1)	7	1*(<80%)	6*( $\geq 80\%$ )
E3 T 4 Procedure Yr 12 Maths	2	(1)	8	2*(<80%)	6*( $\geq 80\%$ )
Pilot 1 T 1 Procedure Yr 13 Art History	4	(1/1/1)	12	12	0
E10 T 1 Procedure Yr 10 Maths	4	(1/1/1)	12	12	0

<sup>1</sup> E = Experiment, Pilot 1 = Pilot Study 1, T = Treatment, Yr 10, Yr 12, Yr 13= Year level at school

<sup>2</sup>. \* These are the unexpected results of the two 2-learning opportunity treatments in Experiment 3. Scores less than 80% in these two treatments counted as confirmations of the Nuthall model whereas scores of 80% or more counted as disconfirmations of the Nuthall model.

on the correct order of the steps for the procedure for creating a tempera painting. While there was some improvement in the performance of the participants overall, the post-test performances of most of the participants were well below what the Nuthall model would have predicted. One possible explanation for the unexpectedly poor post-test performances of the participants in Treatments 2 and 3 is the nature of the procedure they were attempting to recall. Whereas in Treatment 1 participants were being asked to *perform* a procedure,

participants in Treatments 2 and 3 were simply being asked to *describe* a procedure. The results of Treatments 2 and 3 suggest that students may need more learning opportunities to learn and remember how to *describe* a new procedure than they do to *perform* a new procedure.

Table 2 summarises the number of confirmations and disconfirmations of Nuthall's three to four learning opportunity rule for learning and remembering new procedures which have to be performed. Across the six experimental treatments which examined the effect on the retention of procedures of the number of learning opportunities there were 43 confirmations and 12 disconfirmations of Nuthall's rule.

The results of this research provide some additional supporting evidence for the Nuthall model with respect to the number of learning opportunities for the learning and remembering of new procedures where each of the steps can already be performed. It was interesting to note that when students were provided with three or four learning opportunities to learn a procedure that they had to perform every single student succeeded in recalling at least 80 per cent of the steps of the procedures in the correct order. The fact that students in Experiment 3 were able to learn and remember the sequence of steps for graphing both a line and a hyperbola suggests that Nuthall's three to four learning opportunity rule may need to be modified for the learning and remembering of new procedures which can be performed. The results of the present study also suggest not only that the Nuthall model may not apply where the student has to learn how to perform each step as well as learning the correct sequence of steps but also that it may not apply where students only have to describe, rather than perform, a new procedure.

*The inter-study interval between learning opportunities.* The present experiments failed to provide empirical support for the second component of Nuthall's model of the learning and



remembering process, that is, the claim that an inter-study interval between any pair of learning opportunities needs to be no more than two days. The results of Treatment 1 in Experiment 4 where Year 13 Art History participants received four learning opportunities using a 2-day/5-day/2-day schedule of inter-study intervals appeared to provide the first disconfirmation of the Nuthall theory with respect to the critical time interval between pairs of learning opportunities. A 3-day/4-day/3-day schedule of inter-study intervals in Treatment 2 of Experiment 4 also failed to support the 2-day time window.

Table 3

*Number of confirmations and disconfirmations of Nuthall's 2-day time window claim for the recall of unknown knowledge items for the 6 experiments which tested this claim.*

Treatment	No of LOs	Schedule of LOs	N	No of Confirmations	No of Disconfirmations
				< 80%	≥ 80%
E4 T 1 (KI) 13 Art Hist	4	(2/5/2)	13	1	10
E 4 T 2 (KI) 13 Art Hist	4	(3/4/3)	10	0	11
E 5 T 2 (KI) 12 Eng	4	(1/3/1)	14	3	11
E 5 T 3 (KI) 12 Eng	4	(3/3/1)	15	5	10
E 6 T 1 (KI) 13 Eng	4	(3/3/3)	14	0	14
E 6 T 2 (KI) 13 Eng	4	(4/6/7)	14	5	9
E 7 T 1 (KI) 12 Eng	4	(4/3/4)	13	0	13
E 7 T 2 (KI) 12 Eng	4	(6/7/6)	13	4	9
E 8 T 1 (KI) 13 Art Hist	4	(5/6/4)	16	1	15
E 8 T 2 (KI) 13 Art Hist	4	(8/9/5)	16	3	13
E 9 T 1 (KI) 13 Eng	4	(7/7/7)	12	1	11
E 9 T 2 (KI) 13 Eng	4	(14/28/20)	12	7	5

Legend: E = Experiment, T = Treatment, KI = Knowledge Item, 7-13 = Year level at school, Eng = English, Art Hist = Art History

Five further experiments were conducted which examined the effects on retention of extending the length of the inter-study interval beyond Nuthall's two-day time window. All but one of the treatments (Treatment 2 in Experiment 9) in the series of five experiments replicated the results of Experiment 4 using, where possible, different content, different subjects and more challenging schedules of inter-study intervals. As can be seen from Table 3, this lengthy process established that most participants could learn and remember sets of unknown knowledge items even with an inter-study interval between each pair of learning opportunities of up to nine days.

Time constraints and the failure to establish a time window for the learning and remembering of new knowledge items in this series of experiments led to the decision in the final treatment of Experiment 9 to greatly increase the length of the inter-study intervals. This decision was taken to try to see if there was a point at which most participants would fail to learn and remember factual content. It was hoped that if this could be achieved it might be useful to future research pursuing this line of enquiry. The 14-day/28-day/20-day schedule of inter-study intervals used in Treatment 2 of Experiment 9 produced a surprising result. While seven of the 12 participants failed to score at least 80 per cent in their post-tests, the remaining five all scored 85 per cent or more. Three of the five participants scored 100 per cent. These results suggest that for secondary school students the concept of a two-day time window is probably not applicable to the learning and remembering of factual content.

The results of Experiment 10, summarized in Table 4, provided further disconfirmations of Nuthall's two-day rule. This experiment involved the teaching of new procedures for drawing Mathematical graphs to a group of Year 10 participants. In Treatment 2 when participants received four learning opportunities using a 4-day/3-day/3-day schedule of inter-study intervals 12/13 participants scored 91 per cent or more in their post-tests. In Treatment 3 where participants experienced four learning opportunities using either an 11-

day/3-day/6-day or an 11-day/7-day/2-day schedule 7 of 8 students also scored 92 per cent or more.

Table 4

*Number of confirmations and disconfirmations of Nuthall's 2-day time window claim for the recall of unknown procedures in Experiment 10*

Treatment	No of LOs	Schedule of LOs	N	No of Confirmations < 80%	No of Disconfirmations ≥ 80%
E 10 T 2 Procedure Yr 10 Maths	4	(4/3/3)	13	0	12
E 10 T 3 Procedure Yr 10 Maths	4	(11/3/6)	8	1	7

Legend: E = Experiment, T = Treatment, P = Procedure, Yr 10 = Year level at school

There were 126 disconfirmations and 23 confirmations of the Nuthall 2-day rule in the six experiments which examined the effect on the retention of new knowledge items of the length of the inter-study interval between pairs of learning opportunities (See Table 1 in Chapter 13). The results for Treatment 2 of Experiment 9 are not included in this calculation because the very long inter-study intervals used in this treatment were deliberately aimed at trying to prevent students from recalling the knowledge items. There were a further 19 disconfirmations and 1 confirmation of the Nuthall 2-day rule in the two experimental treatments in Experiment 10 which examined the effect on retention of Mathematics procedures of the length of the inter-study interval. Excluding the results of Treatment 2 in Experiment 9, 85.8 per cent of the outcomes results from the seven experiments which tested the validity of Nuthall's 2-day time window hypothesis failed to provide support for the hypothesis.

## **The AB Single-Case Experimental Design**

It needs to be acknowledged that the AB single-case experimental design has significant limitations (Johnston & Pennypacker, 2009). Its essential weakness is that a single experimental demonstration cannot demonstrate whether the changes to the dependent variable are “functionally related to the independent variable” (p. 263). In order to build an argument for a functional relationship between the independent and dependent variables using this experimental design, it is necessary to conduct a series of direct and systematic replications. This was the strategy adopted in the present research.

The AB single-case design was chosen because of the desire to study behaviour changes in individual learners. Barlow et al. (2003) have argued that “until recently...science lacked an adequate methodology for studying behavior change in individuals. This gap in our methodology has retarded the development and evaluation of new procedures in the mental health professions as well as in educational fields” (p. 1). It can be argued that single-case research designs provide the methodology which Barlow et al. claimed had been lacking because “single-case research designs ...make it possible to draw scientifically valid conclusions from the investigation...of individuals” (Blampied, 1999, p. 89).

## **The Reliability of the Experimental Results**

The extent to which a researcher can have confidence in the reliability, that is the reproducibility of his or her observations, is dependent upon the level of internal validity achieved. This can only be determined by replication. The researcher who conducts applied research in real-world classrooms is faced with greater challenges with respect to controlling extraneous variables than the researcher who carries out experiments in a laboratory setting. This makes the achievement of internal validity even more important.

In the present study a high level of treatment fidelity was achieved by having one teacher carry out all ten experiments and the pilot studies. One of the most critical aspects of

the present study which needed to be controlled was the context of both the experiments and pilot studies. In order to create a valid test of Nuthall's model of the learning and remembering process it was imperative that participants did not know that they were participating in an experiment. This was achieved by not informing the participants that they had been involved in an experiment until after their involvement in the study had been completed and by the teacher ensuring that the experimental lessons were conducted as if they were normal classroom lessons.

Marking accuracy was controlled by having one teacher check all the post-tests for each of the eight experiments which involved the learning of new knowledge items and another specialist mathematics teacher check all the post-tests for the two experiments which required participants to learn new Mathematics procedures for drawing graphs. Inter-marker agreement ranged from 99.3 per cent to 100 per cent with an average of 99.7 per cent across the eight knowledge item experiments. For the two procedure experiments in which new procedures were taught, the inter-marker agreement ranged from 97.6 per cent to 98.7 per cent with an average of 98.2 per cent.

One of the greatest threats to the internal validity of experimental findings in a single-case study using an AB design is the possibility of carry-over (sequence) effects. In the present experiments this would have taken the form of pre-test sensitisation, that is, the effect of the pre-test sensitising some of the students to the importance of certain elements of the content of the lessons which followed. In the present experiments this threat to internal validity is likely to have been small since much of the content of each experimental lesson was, in fact, tested by the post-test. It is also the case that the pre-testing procedure was the same for every student in every treatment so any carryover effects were likely to have been similar for every experimental treatment.

Effective controls were exercised over the number of learning opportunities, the set size (the number of knowledge items to be learned), the prior learning of the students and the learning outcomes. Content difficulty was controlled by developing experimental material which was at the appropriate curriculum level for the participants. The activity difficulty was controlled by using a limited number of authentic classroom activities and wherever possible by repeating these from treatment to treatment within each experiment and from experiment to experiment. The number of direct replications within each experimental treatment in each of the ten experiments and the consistent pattern of results they produced suggests that a sufficient degree of control was achieved for the experiments to have a sufficient level of internal validity to allow the drawing of robust conclusions.

It might be argued that a more systematic approach to the scheduling of inter-study intervals, such as using fixed inter-study interval schedules or expanding inter-study intervals, should have been adopted given that there is research to suggest that different types of inter-study schedules can influence learning and remembering in different ways (Roediger & Karpicke, 2011). This was not possible in a school setting. Factors such as timetable constraints, unscheduled interruptions to class time, the demands of the curriculum and the pressures of internal and external assessment create limited windows of opportunity for conducting experiments, and the length of the school term militated against the systematic use of either fixed-interval or expanding schedules of inter-study intervals. Given the data in Table 3 it seems extremely unlikely that the unsystematic scheduling of increasingly lengthy inter-study intervals could account for the results observed.

### **The Generality of the Experimental Results**

In his article entitled *Evidence Based Education Policies: Transforming Educational Practice and Research* Slavin (2000) reminded researchers that “Research in education should ultimately have something to do with improving outcomes for children” (p. 20). This

means that research findings, however they are obtained, must, in the final analysis, be shown to have the same effect in the classroom that they have in the laboratory and they must be shown to have the same effect on individual children that they have on the “average child.” Classroom-based research directly and systematically replicated can achieve such a goal. Classroom-based research which can be directly replicated gives the classroom teacher confidence not only in the reliability of the findings themselves but also in the potential for achieving similar results with different students, while classroom-based research which can be systematically replicated again and again indicates to the classroom teacher that the findings are likely to be generalisable to different contexts and different subject matter.

### **Comparing the Results of the Present Study with those from Previous Studies**

A number of factors make it difficult to compare the results of the present research with previous research. First, only a handful of experimental studies into the effects on retention of distributing learning opportunities in time have been based in real-life classrooms. In addition, very few researchers have sought to control the total number of learning opportunities that experimental participants are exposed to in their research. Thirdly researchers have often been content to observe a statistically significant between groups effect without taking the next step of showing that this effect also occurs at the level of the individual. Finally, the single-case AB experimental design used in the present research is very different to the experimental designs used by most other researchers on learning and memory.

*Previous studies which explore the effects on retention of the number of learning opportunities.* The results of the present research are reasonably consistent with the results of previous studies of the effect on retention of the number of learning opportunities during exposure to authentic, classroom-relevant instruction. (Axelrod, Kramer, Appleton, Rockett,

& Hamlet, 1984; Barbetta, Heward & Bradley, 1993; Cepeda et al., 2009; Childers & Tomasello, 2002; McLay, 2004; McWilliams, 2006; Rawson & Dunlosky, 2011)

*Previous studies on the effects on retention of varying the time window.* The results of the present study failed to identify any time windows for the learning and remembering of either new knowledge items or procedures for learners between 10 and 18 years of age.

*Previous studies of the spacing effect.* The results of the seven experiments in which the effects on retention of varying the inter-study interval between pairs of learning opportunities beyond Nuthall's two-day time window were examined appear to be quite consistent with the group effects reported by Cepeda et al. (2008) and Cepeda et al. (2009). It will be recalled that Cepeda et al. (2008) examined the effect on both recognition and recall of varying the inter-study intervals between two separate learning sessions. While Cepeda et al. did not strictly control the number of learning opportunities in the first learning session, the total number of learning opportunities subjects received for each new trivia fact was likely to have been quite similar to the four learning opportunities provided in the seven experiments which varied the inter-study interval during the present research. With a retention interval of 7 days, the subjects in Cepeda et al.'s study recalled at least 90 per cent on average of the new trivia facts they were required to learn and did so when their two learning sessions were separated by inter-study intervals of 0-, 1-, 2-, 7- and 21 days. Even with a 105-day inter-study interval, students were still able to recall over 80 per cent on average of the new trivia facts.

The results of the present study are also consistent with the findings of Cepeda et al. (2009). It needs to be noted that in their first experiment in which students received two learning sessions to learn and remember 40 Swahili-English word pairs, the number of learning opportunities in the first learning session was, as in Cepeda et al.'s (2008) experiment, not strictly controlled. With a retention interval of 10 days students recalled



about 70 per cent of the Swahili-English word pairs after two learning sessions regardless of whether those sessions were separated by inter-study intervals of 1-, 2-, 4-, 7- or 14 days.

It needs to be emphasized, however, that the present study was not a spacing effect study. Moreover, the fact that this was not a parametric study means that its results can add little to the debate about such questions as how to optimize retention by distributing practice over time or how the relationship between the length of the inter-study interval and the length of the retention interval affects retention. The most advanced treatment of these important questions is to be found in the work of Cepeda et al. (2008) and Cepeda et al. (2009).

### **Theoretical Implications of the Results of the Present Research**

*Theories of working memory.* Working memory remains “one of ‘hottest’ topics in cognitive psychology and cognitive neuroscience” (Miyake & Shah, 1999, p. xiii). While today researchers have at their disposal highly complex brain scanning techniques, Ruchkin et al. (2003) make the important observation that, despite these technological advances and the confidence some researchers have in their models of memory, all researchers in the field of brain research are, at best, using “blunt instruments” (p. 764) whose combined capacity still remains far from capable of capturing all of “the degrees of freedom” (p. 764) that our complex brains possess.

Whereas Nuthall felt compelled to develop his own model of long-term working memory to accommodate the findings of his descriptive studies, the findings of the present study can be readily accommodated by a number of the current main models of working memory (e.g. Baddeley, 2007; Cowan, 2005; Ruchkin et al., 2003). According to Nuthall’s model, three to four representations of a new knowledge item needed to come together in working memory for a new knowledge construct to form and then be transferred to long-term memory. In the present study it was clear that representations of some new knowledge items were being transferred to long-term memory even after just two learning opportunities. It is

important to acknowledge that the very real possibility exists that some new knowledge items might have been able to be transferred to long-term memory after a single learning opportunity. Consequently the results of the present research can be accommodated by any model of working memory which allows for the possibility of the transfer of a representation of a new knowledge item into long-term memory even after a single learning opportunity. Ruchkin et al.'s (2003) model, for example, allows for this possibility. They argue that even when confronted with a new fact or proposition that new knowledge will still be associated with something in a person's long-term memory because "human thought simply cannot exist independent of already known memory elements" (Ruchkin et al., 2003, p. 768).

*The time window construct.* The results of the present research can add little to the debate about the validity of the time window construct as defined by Rovee-Collier (1995). The present study was only able to establish that for students between the ages of 10 to 18 Nuthall's critical two-day gap between learning opportunities was not a time window for the learning and remembering either of knowledge items or learning the sequence of steps in a procedure that must be performed. The failure in the present research to find a time-window for the learning and remembering of either knowledge items or the sequence of steps for a specific type of procedure in this age group should not be seen as throwing doubt on the validity of the time window construct. Rovee-Collier's (1995) conception of the nature of the time window construct anticipated the data collected in this research. In her 1995 article she readily acknowledged that time windows were likely to be more difficult to determine in older individuals than in infants:

Consider...that (a) every retrieval from memory affects the width of a given time window, (b) memory retrieval can be either direct or indirect and may be neither overtly expressed...nor even be accompanied by awareness...and (c) individuals have different, complex networks of associations by which prior information can be accessed and retrieved. Moreover, the opportunity for all of these factors to operate increases with age-related experience (Rovee-Collier, 1995, p. 148).

As a consequence of the factors identified above, Rovee-Collier argued that differences in people's life experiences were very likely to result in people developing very different time windows.

*Cognitive and neurobiological theories of the spacing effect.* While the data collected in this research were not designed to specifically address the validity of current theories of the spacing effect, the data are consistent with some of the main theories of the spacing effect and with recent neurobiological explanations as to why some memories persist longer than others (Bekinschtein et al., 2007; Medina et al., 2008). The present research findings are consistent with consolidation theories and encoding variability theories but not with deficient processing and study-phase retrieval theories. Deficient processing theories may serve to explain why there may be differences in the recall of massed versus spaced items but they do not appear to offer an effective explanation for the results achieved in the present study given that, in all treatments but one, every pair of learning opportunities was separated by at least a day. The results of the present study do not appear to provide any data relevant to study-phase retrieval theories because the retention interval was kept at one week for all the experiments.

In the present experiments, the fact that three or four learning opportunities distributed in time produced such high levels of recall among students of different levels of academic ability provides support for the view that basic neural machinery which demands the spaced repetition of learning experiences is involved for the learning and remembering of certain types of knowledge. As suggested previously, if a theta train is indeed the equivalent of a learning opportunity and if maximal long-term potentiation represents the formation of a long-term memory then the finding of Kramar et al. (2012) that three theta trains, each separated by an hour, are required to produce maximal potentiation hints faintly at the

possibility that for mammalian brains a specific number of learning experiences distributed in time is required for the formation of memories.

Ultimately, however, we may discover that memory formation is, as Zull (2011) argues, basically a question of statistical probability:

The behavior of existing synapses in the moment depends on their past behavior. This suggests that neuronal networks function as statistical systems. Evolution has given us biochemical processes for keeping strong networks strong and weak ones weak...Just like evolution, memory is not driven by need or by design, but rather by probability and selection. What looks like instruction is just selection. Many possibilities exist in the immense collection of the networks of the brain, but only the useful ones, the ones that fire a lot, are statistically most likely to fire again...It is the experience of each neuron in a pattern that guides memory formation (p.185).

### **Implications for Teachers**

Nuthall's analysis of the results of his six descriptive studies led him to question long-held assumptions about the role which general ability plays in student learning. In his studies Nuthall found that if students experienced the complete set of information needed to learn and remember a fact or concept on three or four occasions then they would learn and remember that fact or concept regardless of their level of ability. This led him to argue that more and less able students learn in essentially the same way. Next, on the basis of his observations, Nuthall argued that certain students learned and remembered more in his studies not because of their general ability but because of their skill in generating additional learning opportunities in the relatively unstructured environment of Social Studies and Science lessons in the primary school classroom. Thus, on the basis of his research, Nuthall claimed that "individual differences that emerge in students' achievement in school are the product of individual differences in the ways in which the students participate in, and internalise, classroom learning activities" (Nuthall, 1999b, p. 213).

Confirmation of Nuthall's claim that students need a number of learning opportunities in order to learn and remember new knowledge items has important implications for teachers.

The results of the experiments reported herein suggest that all students, regardless of ability, are capable of learning and remembering new knowledge if they experience a sufficient number of learning opportunities. It follows that teachers play a pivotal role in determining what students will and will not learn and remember in their classrooms. In particular, the present research suggests that teachers need to provide all students with a sufficient number of learning opportunities to enable them to learn and to remember essential curriculum knowledge. The present study also demonstrated the importance of having regular tests of student learning and remembering. Not only does the test provide an additional learning opportunity but it also identifies those students who may need extra learning opportunities to learn and to remember important new knowledge. The evidence presented here strongly suggests that the teacher can significantly influence what students learn and remember in a new teaching unit. It follows that when a student fails to learn a new knowledge item it may be prudent to look first at one's own teaching practices and not the academic ability of the student. This will require a paradigm shift for many teachers. Teachers who are prepared to accept more responsibility for the learning of their students may appreciate that a corollary of this may be a more positive view of student potential than that suggested by standardised tests of academic ability.

The findings of the present research also suggest that teachers have an important role to play in teaching students how to generate their own learning opportunities. Nuthall's studies and the findings of the present research hint at the possibility that differences in student learning outcomes may be the result of differences in students' ability to generate their own learning opportunities, rather than the result of differences in ability. One of the first metacognitive strategy lessons might involve teaching students the critical importance of having a sufficient number of learning opportunities when attempting to learn not only new

procedures but also new facts and concepts. A second lesson might focus on teaching students to develop self-rehearsal strategies.

The present studies failed to confirm Nuthall's two-day rule and found that inter-study intervals between pairs of learning opportunities of up to nine days for the recall of knowledge items and up to 11 days for the recall of procedures did not impair students' recall. The results indicate that teachers have a great deal of flexibility when it comes to scheduling learning opportunities. It will be recalled that all of the experiments in this study took place against a background of unprecedented seismic activity. In spite of this, most participants continued to score at least 80 per cent or more in their individual post-tests when faced with lengthy inter-study intervals between learning opportunities. This is clear evidence of the robust nature of the spacing effect.

The present research helps to "illuminate the links between what the teacher does and what the student learns" (Nuthall & Alton-Lee, 1998, p. 25). It strongly suggests that not only the timing and frequency of the learning opportunities but also the very nature of the learning opportunities themselves are of critical importance in determining what will be learned and remembered. Ideally for the teaching of new knowledge items, learning opportunities should be three term learning opportunities as these provide students with the complete set of information they need to learn and to remember them. It is also desirable that the activities which provide the learning opportunities be sufficiently varied and challenging to require students to engage with the to-be-learned content in a variety of ways. Like Nuthall (1999a), neuroscience suggests that it is the nature of the "learning experiences", in addition to their timing and frequency, which is important in determining what is learned and remembered. The magnitude of the theta train bursts are a critical factor in determining the duration of LTP (Zull, 2011). If LTP is an analogue of long-term memory then perhaps biology explains why, as Nuthall suggested, the activities which a teacher chooses to assist students to learn and

remember new knowledge cannot involve just simple repetition (Nuthall, 1999a). Given that LTP “is triggered by intense and repeated neuron firing” (Zull, 2011, p. 185), it follows, as Nuthall (1999a) has suggested, that for the formation of long-term memory students need a variety of tasks which pose “desirable difficulties” (Bjork, 1994) and demand of them an “effort from within” (James, 1981).

One of the most important implications for teachers from not only the present research but also Nuthall’s studies is that “students’ memories are structured by school activities, and the extent to which...[the students] remember experience is a product of that structuring” (Nuthall & Alton-Lee, 1997, p. 7). As Alton-Lee (2006) suggests, by managing the timing and the number of learning opportunities and by designing effective learning activities, the teacher creates “the very building blocks of students’ memories” (p. 621).

The present research also has important implications for the teaching of procedures. During a pilot study, a class of Year 12 Mathematics students failed to learn and to remember the 12 to 14 step procedures for drawing a circle graph when provided with four learning opportunities. It became apparent after interviewing some of the participants and analysing their post-test performances that many of the participants had failed to complete the procedure because they did not know how to perform one or more of the key steps in the procedure. This gave rise to the realisation that the pilot study was not a true test of the Nuthall model for the teaching of procedures because it provided four learning opportunities not only to learn and remember a new 12 to 14 step procedure but also to learn and remember how to perform each of the 12 to 14 steps in the procedure. With this came the further realisation that in order to carry out a true test of the Nuthall model for the teaching of procedures it was imperative that the participants know how to perform most of the essential steps in the procedure prior to teaching the actual sequence of steps in the procedure. When this was done, a group of Year 12 Mathematics participants in Experiment 3 demonstrated

that, in addition to completing a parabola and a circle graph with four and three learning opportunities respectively, they could even acquire the procedure for drawing a line graph and a hyperbola with just two learning opportunities. These experiments revealed that the 3 to 4 learning opportunity rule may not apply to all learning outcomes. The results of the present research suggest that students may need more than four learning opportunities not only to learn and remember procedural descriptions (without performance of the procedure) but also to learn both the steps of a new procedure and the sequence of steps for that procedure. The Mathematical procedures experiments also revealed that for the effective teaching of a procedure a critical preliminary step is to make certain that students have learned and remembered how to complete each of the steps in the procedure.

The findings of the present research add support to a growing number of investigators who advocate changes to way in which the curriculum is delivered in light of the empirical evidence that distributing learning opportunities in time is required for student learning and remembering. Given that three to four learning opportunities distributed in time produces such a high level of recall in all students not only of new sets of knowledge items but also of new procedures and operations, the curriculum and educational textbooks might better serve student interests if they were structured so that essential facts, skills and ideas were deliberately and systematically revisited during the course of the academic year.

### **Implications for Research on Teaching**

The present research demonstrates that controlled experimental learning research can be undertaken in real world classrooms. Much of the research conducted in classrooms today is qualitative or descriptive research and while it serves many useful purposes “it cannot be used to answer questions which require an experimental analysis... [such as] questions about how learning occurs” (Church, 1999c, p. 3). To answer such questions it is necessary to conduct an experimental study. There are many examples in the literature of experimental



studies being conducted successfully in classrooms (e.g. Gardner, Heward & Grossi, 1994; Davis & O'Neill, 2004; Narayan, Heward, Gardner, Courson, & Omness, 1990; Kubina, Amato, Schwilk, & Therrien, 2008). The present study provides a further demonstration that it is possible to conduct learning experiments in classrooms with authentic curriculum content provided that great care is taken to control all the confounding variables. The great advantage for teachers of the single-case experimental method used in this study is that it allows them to study the learning of individuals. "The changes which we refer to as learning are changes which are occurring at the level of the individual. Any procedure we use to study learning must, therefore, be a procedure which is capable of tracking changes in the motivation, competency and attitudes of individual learners" (Church, 1999c, p. 10).

### **The Trials and Tribulations of Doing Classroom-Based Research**

Researchers seeking to undertake classroom-based research may well benefit from talking with teachers about the potential for disruptions to an "average" teaching day. Schools are busy places and unexpected interruptions to a scheduled class period can come at the most inopportune of times. In addition to interruptions caused by such things as assemblies running over time and students being out of class due to bouts of illness or through their participation in extra-curricular activities for which no prior warning was given, Mother Nature can intervene in surprising ways. The students who participated in these experiments, for example, have had to endure over 14,000 earthquakes since 4 September 2010. However, despite the trials and tribulations of classroom research, it can be argued that only in the crucible of the classroom can the real worth of educational research ultimately be determined.

### **Future Research**

The first possible area of research relates to the search for the most effective means of promoting "durable and efficient learning" (Rawson & Dunlosky, 2011, p. 300). Nuthall's

model was based to a large extent on the post-test results the students in his descriptive studies achieved with a retention interval of a year. In the present experiments the retention interval remained constant at one week. One possible area of research in the future would be to investigate in a parametric experimental study the durability of the learning which results when students receive three to four learning opportunities to learn and remember a new set of knowledge items.

The work of Cepeda et al. (2009) suggests another area of possible research. Their research strongly suggests that there may be a relationship between the length of the retention interval and the most effective inter-study intervals. It will be recalled that Cepeda et al. (2009) demonstrated a *non-monotonic* relationship between the inter-study interval and the retention interval so that up to an optimal point, an increase in the inter-study interval increases the retention interval. If these findings could be replicated using just three to four learning opportunities it provides teachers with plenty of opportunity to design schedules of learning opportunities which enable students to perform at their best when they are faced with sitting external examinations which are some months away.

Mozer et al. (2009) have quite recently developed their Mutliscale Context Model (the MCM model) for predicting the effect on retention of different learning material of using a particular schedule of inter-study intervals. Mozer et al. tested the predictive power of their model by comparing the actual results of a series of experiments which examined the effect on retention of the relationship between the size of the inter-study interval and the length of the retention interval with the results that their model generated for the same experiments. The model's predictions proved to be extremely accurate. Thus another possible area of future research might be to explore the extent to which the results of the research proposed above are consistent with the predictive power of the Mozer model.

Another possibility for future research is to investigate how the temporal distribution of three to four learning opportunities affects different learning outcomes. A future study might explore the effect on the learning and remembering of new names and definitions in subjects such as Chemistry and Physics of three to four spaced learning opportunities. Alternatively, following the suggestion of Delaney et al. (2010), investigators might wish to explore how the spacing of three to four learning opportunities affects the learning and remembering of more complex skills such as the learning of new spelling words and new critical thinking strategies.

Finally, another area of possible future research suggested by the present experiments is the systematic investigation of Nuthall's questioning of the role that academic ability plays in both how we learn and what we learn. Nuthall's bold claims are very much at odds with a great body of research which demonstrates that academic ability is strongly related to student learning in the classroom (Carroll, 1993). If Nuthall's claims are valid, then the relationship between academic ability and academic achievement may be demystified and this, in turn, may force some teachers to re-examine their expectations with respect to the learning abilities of the low achieving students in their classrooms.

## **Conclusion**

The experiments described in this thesis tested the two key components of the Nuthall model of the learning and remembering process. The first component of the Nuthall model states that students require three to four learning opportunities in order to learn and remember new knowledge items. The results of Experiments 1 and 2 were consistent with the first component of the Nuthall model, that is, the number of learning opportunities students require to learn and remember new knowledge items. However, the results of Experiment 3 did not confirm the first component of the Nuthall model for the learning of new procedures.

Most of the Year 12 Mathematics students in this experiment required only two learning opportunities in order to learn and remember two new mathematical operations.

The second component of the Nuthall model states that to learn and remember new knowledge the inter-study intervals between the learning opportunities must not exceed two days. The first experimental test of the timing component of the model (during the teaching and practice of new knowledge items) failed to provide supporting evidence of the two-day rule. Five replications of this experiment using different subjects, different content and different schedules of inter-study intervals provided five further disconfirmations of the two-day rule. It was found that inter-study intervals of up to eight and nine days between pairs of learning opportunities did not prevent most students from achieving very high rates of recall of previously unknown facts in their individual post-tests. The results of Experiment 10 which required Year 10 Mathematics participants to learn and remember new procedures for drawing graphs provided a further disconfirmation of the two-day rule. Similar results have been reported by others such as Cepeda et al. (2008) and Cepeda et al. (2009).

The value of employing a single-case experimental study in a real life classroom with authentic curriculum content to test the predictive validity of Nuthall's model of the learning and remembering process was clearly demonstrated. This experimental methodology provided an accurate and detailed picture of the learning of individual participants in each of the experiments in the present study.

The principal cognitive theories of working memory, the time window construct and the spacing effect and recent neuroscientific hypotheses in these areas were outlined. It was found that the results of the present study were consistent with some of the main theories of both working memory and the spacing effect. While the results of the present study clearly suggested that Nuthall's 2-day time rule was not a time window they could not be used to draw any conclusions about the validity of the time window construct. The neurobiological

research, although extremely interesting, has yet to identify the specific neural mechanisms which result in remembering in humans.

The findings of the present study strongly suggest that, in order to assist students to learn and remember essential curriculum knowledge, teachers might provide four learning opportunities with inter-study intervals of up to nine days between pairs of learning opportunities. The present findings also suggest that teachers should be encouraged to take care with the construction or selection of activities which provide the essential learning opportunities. It would appear that activities which are challenging and force students to engage with the to-be-learned material provide more memorable learning opportunities. Finally, the results of this study provide yet more evidence of the wisdom of designing curricula and textbooks which enable students to revisit essential curriculum knowledge on several occasions during the course of the academic year.

It was suggested that future research could investigate the relationship between the length of the inter-study interval and the length of the length of the retention interval, the extent to which the results of that proposed research correlated with the predictive power of Mozer et al.'s Multiscale Context Model, the impact of three or four learning opportunities distributed in time on the performance of more complex critical thinking and comprehension skills, and the role that prior achievement and differences in academic ability play in creating differences in student learning outcomes during well-designed instruction with adequate levels of repetition.

## REFERENCES

- Abraham, W. C., Logan, B., Greenwood, J. M., & Dragunow, M. (2002). Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. *The Journal of Neuroscience*, 22, 9626-9634.
- Albensi, B. C., Oliver, P. R., Taupin, J., & Otero, G. (2007). Electrical stimulation protocols for hippocampal synaptic plasticity and neuronal hyper-excitability. *Experimental Neurology*, 204, 1-13.
- Albers, A. E., & Greer, R. D. (1991). Is the three-term contingency trial a predictor of effective instruction? *Journal of Behavioral Education*, 1, 337-354.
- Alton-Lee, A. G. (1984). *Understanding learning and teaching: An investigation of pupil experience and content in relation to immediate and long-term learning*. (Unpublished doctoral dissertation). University of Canterbury, Christchurch, New Zealand.
- Alton-Lee, A. G., & Nuthall, G. A. (1990). Pupil experiences and pupil learning in the elementary classroom: An illustration of a generative methodology. *Teaching and Teacher Education*, 6, 27-45.
- Alton-Lee, A. G., & Nuthall, G. A. (1992, April). *Students learning in classrooms: Curricular, instructional and sociocultural processes influencing student interaction with curriculum content*. Annual Meeting of the American Educational Research Association, San Francisco, USA.
- Alton-Lee, A. G., Nuthall, G. A., & Patrick, J. (1993). Reframing classroom research: A lesson from the private world of children. *Harvard Educational Review*, 63, 50-84.
- Alton-Lee, A. G. (2006). How teaching influences learning: Implications for educational researchers, teachers, teacher educators and policy makers. *Teaching and Teacher Education*, 22, 612-626.
- Ary, D., Jacobs, L. C., & Sorensen, C. (2010). *Introduction to research in education* (8<sup>th</sup> ed.). Belmont, CA: Wadsworth, Cengage Learning.
- Atkinson, R. C., & Shrifin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. P. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York, NY: Academic Press.
- Ausubel, D. P. (1966). Early versus delayed review in meaningful learning. *Psychology in the Schools*, 3, 195-198.
- Axelrod, S., Kramer, A., Appleton, E., Rockett, T., & Hamlet, C. (1984). An analysis of the relevance of topographical similarity on positive practice of spelling errors. *Child and Family Behaviour Therapy*, 6, 19-32.

- Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, 302, 311-324.
- Baddeley, A. D. (2007). *Working memory, thought and action*. New York, NY: Oxford University Press.
- Baddeley, A.D., & Hitch, G. J. (1974). Working memory. In G. J. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-89). New York : Academic Press.
- Barbetta, P. M., Heward, W.L., & Bradley, D.M.C. (1993). Relative effects of whole word and phonetic-prompt error correction on the acquisition and maintenance of sight words by students with developmental difficulties. *Journal of Applied Behavior Analysis*, 26, 99-100.
- Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behaviour change* (3<sup>rd</sup> ed.). Boston, MA: Pearson Education Ltd.
- Bekinschtein, P., Cammarota, M., Igaz, L. M., Bevilacqua, L. R. M., Izquierdo, I., & Medina, J. H. (2007). Persistence of long-term memory storage requires a late protein synthesis and BDNF-dependent phase in the hippocampus. *Neuron*, 53, 261-277.
- Bekinschtein, P., Cammarota, M., Katche, C., Slipczuk, L., Rossato, T. I., Goldin, A., Izquierdo, I., and Medina, J. H. (2008). BDNF is essential to promote persistence of long-term memory storage. *Proceedings of the National Academy of Sciences of the United States*, 105, 2711-2716.
- Blampied, N. M. (1999). A legacy neglected: Restating the case for single-case research in cognitive behaviour therapy. *Behaviour Change*, 16, 89-104.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185-205). Cambridge, MA: MIT Press.
- Bliss, T. V. P., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation in the perforant path. *Journal of Physiology*, 232, 331-356.
- Butler, A. C., Roediger, H. L., & Karpicke, J. D. (2008). Correcting a meta-cognitive error: Feedback enhances retention of low confidence correct responses. *Journal of Experimental Psychology. Learning, Memory and Cognition*, 34, 918-928.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analysis studies*. New York, NY: Cambridge University Press.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171-246). Chicago: Rand McNally.

- Cepeda, N. J., Pashler, H., Rohrer, D., Vul, E., & Wixted, J. T. (2006). Distributing practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354-380.
- Cepeda, N. J., Pashler, H., Rohrer, D., Vul, E., and Wixted, J. T. (2008) Optimizing retention using distributed practice: Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, 19, 1095-1102. doi: 10.1111/j.1467-9280.2008.02209.x.
- Cepeda, N. J., Coburn, N., Mozer, M. C., Pashler, H., Rohrer, D., and Wixted, J. T., (2009) Optimising distributed practice: Theoretical analysis and practical implications. *Experimental Psychology*, 56, 236-248. doi: 10.1027/1618-3169.56.4.236.
- Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, 38, 967-978.
- Church, R. J. (1999a). *Basic learning processes*. Christchurch, NZ: University of Canterbury Department of Education.
- Church, R. J. (1999b). *Instructional processes*. Christchurch, NZ: University of Canterbury Department of Education.
- Church, R. J. (1999c, November). The relative utility of qualitative, social science, and natural science research into learning and teaching. Paper presented to the AARE-NZARE Annual Conference, Melbourne, Australia.
- Cohen, J. (1994). The earth is round ( $p < .05$ ). *American Psychologist*, 49, 997-1003.
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education*. (7<sup>th</sup> ed.). New York, NY: Routledge.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied Behavior Analysis*. Columbus, OH: Merrill Publishing Co.
- Cowan, N. (2005). *Working memory capacity: Essays in cognitive psychology*. New York, NY: Psychology Press.
- Cresswell, J. W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4<sup>th</sup> ed.). Boston, MA: Pearson.
- Davis, L. L., & O'Neill, R. E. (2004). Use of response cards with a group of students with learning disabilities including those for whom English is a second language. *Journal of Applied Behavior Analysis*, 37, 219-222.



- Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation*, 53, 63-148.
- Delaquadri, J. C., Greenwood, C. R., Stretton, K., & Hall, R. V. (1983). The peer tutoring spelling game: A classroom procedure for increasing opportunity to respond and spelling performance. *Education and Treatment of Children*, 6, 225-239.
- Dochy, F., Segers, M., Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, 69, 145-186.
- DeZazzo, J., & Tully, T. (1995). Dissection of memory formation: From behavioral pharmacology to molecular genetics. *Trends in Neurosciences*, 18, 212-218.
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. New York: Teachers' College, Columbia University (Original work published 1885, translated 1913).
- Englemann, S., & Carmine, D. (1991). *Theory of instruction: Principles and applications* (Rev. ed.). Eugene, OR: ADI Press.
- Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.
- Eckel-Mahan, K. L., Phan, T., Han, S., Wong, H., Chan, G. C-K., Scheiner, Z. S., Storm, D. R. (2008). Circadian oscillation of hippocampal MAPK activity and cAMP: Implications for memory persistence. *Nature Neuroscience*, 11, 1074-1082.
- Fisher, R. A. (1935) The design of experiments. In J.H. Bennett (Ed.), *Statistical methods, experimental design, and scientific inference: A reissue of statistical methods for research workers, the design of experiments, and statistical methods and scientific inference*. New York: Oxford University Press.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2010). *Applying educational research* (6<sup>th</sup> ed.). Boston, MA: Pearson.
- Gagne, R. M., & Dick, W. (1983). Instructional psychology. *Annual Review of Psychology*, 34, 261-295.
- Gagne, R. M. (1985). *The conditions of learning* (4<sup>th</sup> ed.). Orlando, FL: Holt, Rinehart and Winston.
- Gardner, R., Heward, W. L., & Grossi, T. A. (1994). Effects of response cards on student participation and academic achievement: A systematic replication with inner-city students during whole-class science instruction. *Journal of Applied Behavior Analysis*, 27, 63-71.

- Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and applications*. (9<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson.
- Glenberg, A. M., & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory and Cognition*, 8, 525-538.
- Gong, L., He, L., Dong, Z., Lu, X., Poo, M., & Zhang, X. (2011). Postinduction requirement of NMDA receptor activation for late-phase long-term potentiation of developing retinotectal synapses in vivo. *The Journal of Neuroscience*, 31, 3328-3335.
- Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1984). Opportunity to respond and academic performance. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education*. Columbus, OH: Charles E. Merrill.
- Greer, R. D. (1994). The measure of a teacher. In R. Gardner, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. W. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 161-172). Pacific Grove, CA: Brooks/Cole.
- Greer, R. D. (2002). *Designing teaching strategies: An applied behavior analysis systems approach*. Orlando, Florida: Academic Press.
- Hattie, J. (2012). *Visible learning for teachers: Maximising impact on learning*. London, England: Routledge.
- Hattie, J., Biggs, J., & Purdue, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66, 99-136.
- Hayne, H. (1990). The effect of multiple reminders on long-term retention in human infants. *Developmental Psychobiology*, 23, 453-477.
- Heward, W. L. (1994). Three "low-tech" strategies for increasing the frequency of active student response during group instruction. In R. Gardner, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. W. Eshleman & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 283-320). Pacific Grove, CA: Brooks/Cole.
- Hintzman, D. L., & Block, R. A. (1973). Memory for the spacing of repetitions. *Journal of Experimental Psychology*, 99, 70-74.
- Hsu, V. C. (2010). Time windows in retention over the first year-and-a-half of life: Spacing effects. *Developmental Psychobiology*, 52, 764-774. doi: 10.1002/dev.20472
- Huang, Y. Y., & Kandel, E. R. (2004). Recruitment of long-lasting and protein kinase A-dependent long-term potentiation in the CA1 region of the hippocampus requires repeated tetanization. *Learning and Memory*, 1, 74-82.
- Izquierdo, I., Bevilacqua, L. R. M., Rossato, J. I., Bonini, J. S., Medina, J. H., & Cammarota, M. (2006). *Trends in Neurosciences*, 29, 496-505.

- Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17, 649-667.
- James, W. (1981). *The principles of psychology* (Vol 1). Cambridge, M.A: Harvard University Press (Original work published 1890).
- Johnston, J. M., & Pennypacker, H. S. (1993). *Readings for strategies and tactics of behavioral research* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kameenui, E. J., & Simmons, D. C. (1990). *Designing instructional strategies: The prevention of academic learning problems*. Columbus, OH: Merrill Publishing.
- Kornel, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, 25, 498-503.
- Kornmeier, J., & Sasic-Vasic, Z. (2012). Parallels between spacing effects during behavioral and cellular learning. *Frontiers in Human Neuroscience*, 6, 1-5. doi: 10.3389/fnhum.2012.00203
- Kramar, E. A., Babayan, A. H., Gavin, C. F., Cox, C. D., Jafari, M., Gall, C. M., Rumbaugh, G., & Lynch, G. (2012). Synaptic evidence for the efficacy of spaced learning. *Proceedings of the National Academy of Sciences*, 109, 5121-5126. doi: 10.1073/pnas.1120700109
- Kubina Jr., R. M., Amato, J., Schwillk, C. L., & Therrien, W. J. (2008). Comparing performance standards on the retention of words read correctly per minute. *Journal of Behavioral Education*, 17, 328-338.
- Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review*, 76, 82-96.
- Lynch, G., Rex, C. S., Chen, L. Y., & Gall, C. M. (2008). The substrates of memory: Defects, treatments and enhancement. *European Journal of Pharmacology*, 585, 2-13. doi: 10.1016/j.ejphar.2007.11.082
- Lynch, G., Kramar, E. A., Babayan, A. H., Rumbaugh, G., & Gall, C. M. (2013) Differences between plasticity thresholds result in new timing rules for maximizing long-term potentiation. *Neuropharmacology*, 64, 27-36. [http:// dx.doi.org/10.1016/j.neuropharm.2012.07.006](http://dx.doi.org/10.1016/j.neuropharm.2012.07.006)
- Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Neurosciences*, 11, 442-450.
- McDonald, F. J. (1959). *Educational psychology*. San Francisco, USA: Wadsworth Publishing Company.
- McGaugh, J. L. (2000). Memory - a century of consolidation. *Science*, 287, 248-250.

- McLay, L. K. (2004). *Acquisition, retention and generalisation of object names in 4 year old children during child initiated and adult initiated learning interaction*. (Unpublished MEd Dissertation). University of Canterbury, Christchurch, New Zealand.
- McWilliams, K. G. (2006). *An analysis of variables affecting instructional efficiency*. (Unpublished doctoral dissertation). University of Canterbury, Christchurch, New Zealand.
- Medina, J. H., Bekinschtein, P., Cammarota, M., & Izquierdo, I. (2008). Do memories consolidate to persist or do they persist to consolidate? *Behavioural Brain Research*, 192, 61-69.
- Melton, A. W. (1967). Repetition and retrieval from memory. *Science*, 158, 532.
- Menzel, R., Manz, G., Menzel, R., & Greggers, U. (2001). Massed and spaced learning in honeybees: The role of CS, US, the intertribal interval, and the test interval. *Learning and Memory*, 8, 198-208.
- Mertens, D. M. (2005). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage Publications Inc.
- Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.
- Mozer, M. C., Pashler, H., Cepeda, N., Lindsey, R., & Vul, E. (2009). Predicting the optimal spacing of study: A multiscale context model of memory. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol 22, pp. 1321 – 1329). La Jolla, CA: NIPS Foundation.
- Muzzio, I. (1994). *The effects of timing of postevent information on infant retention*. (Unpublished Master's Dissertation), Rutgers University, New Jersey, United States of America.
- Naqib, F., Sossin, W. S., & Farah, C. A. (2012). Molecular determinants of the spacing effect. *Neural Plasticity*, 2012, 1-8. doi: 10.1155/2012/581291
- Narayan, J. S., Heward, W. L., Gardner III, R., Courson, F. H., & Omness, C. K. (1990). Using response cards to increase student participation in an elementary classroom. *Journal of Applied Behavior Analysis*, 23, 483-490.
- Nuthall, G. A. (1996, December). *What role does ability play in classroom learning?* Annual meeting of the New Zealand Association for Research in Education, Nelson, New Zealand.
- Nuthall, G. (1999a). The way students learn: Acquiring knowledge from an integrated science and social studies unit. *The Elementary School Journal*, 99, 303-341.

- Nuthall, G. A. (1999b). Learning how to learn: the evolution of students' minds through the social processes and culture of the classroom. *International Journal of Educational Research*, 31, 141-256.
- Nuthall, G. A. (2000a). The role of memory in the acquisition and retention of knowledge in science and social studies units. *Cognition and Instruction*, 18, 83-139.
- Nuthall, G. A. (2000b). The anatomy of memory in the classroom: Understanding how students acquire memory processes from classroom activities in science and social studies units. *American Educational Research Journal*, 37, 247-38
- Nuthall, G. A., (2001). Procedures for identifying the information content of student classroom experiences and predicting student learning. Retrieved from the Nuthall Trust website: [http:// www.nuthalltrust.org.nz](http://www.nuthalltrust.org.nz)
- Nuthall, G. A., (2005). The cultural myths and realities of classroom teaching and learning: A personal journey. *Teachers College Record*, 107, 895-934.
- Nuthall, G. A. (2007). *The hidden lives of learners*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Nuthall, G. A., & Alton-Lee, A. G. (1991). *Understanding learning and teaching project*. Phase Two, Final Report to the Ministry of Education. Christchurch, NZ: University of Canterbury Education Department.
- Nuthall, G. A., & Alton-Lee, A. G. (1993). Predicting learning from student experience of teaching: A theory of student knowledge construction in classrooms. *American Educational Research Journal*, 30, 799-840.
- Nuthall, G. A., & Alton-Lee, A. G. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal*, 32, 185-223.
- Nuthall, G. A., & Alton-Lee, A. G. (1997). *Understanding learning in the classroom*. Report to the Ministry of Education Understanding Learning and Teaching Project 3. Wellington, NZ: Ministry of Education.
- Nuthall, G. A., & Alton-Lee, A. G. (1998). *Inclusive instructional design: Theoretical principles emerging from the understanding learning and teaching project*. Report to the Ministry of Education Understanding Learning and Teaching Project 3. Wellington, NZ: Ministry of Education.
- O'Toole, V. M. E. (2005). *Thank you for giving me the opportunity to actually tell you what I know: The role of emotion in children's learning task engagement in the elementary school classroom*. (Unpublished doctoral dissertation). University of Canterbury, Christchurch, New Zealand.

- Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31, 3-8.
- Pressley, M., Levin, J. R., & Mc Daniel, M. A. (1987). Remembering versus inferring what a word means. Mnemonic and contextual approaches. In M. G. Mc Keown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Randolph, J. J., (2007). Meta-analysis of the research on response cards: Effects on test achievement, quiz achievement, participation, and off-task behavior. *Journal of Positive Behavior Interventions*, 9, 113-128.
- Rasch, B., & Born, J. (2007). Maintaining memories by reactivation. *Current Opinion in Neurobiology*, 17, 698-703.
- Rawson, K. A., & Dunlosky, J. (2011). Optimising schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology*, 3, 283-302.
- Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181-210. doi: 10.1111/j.1745-6916.2006.00012.x
- Rohrer, D. & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology*, 20, 1209-1224.
- Rose, S. (1992). *The making of memory: From molecules to mind*. New York, NY: Doubleday.
- Rovee-Collier, C. (1995). Time windows in cognitive development. *Developmental Psychology*, 31, 147-169.
- Rovee-Collier, C., Greco-Vigorito, C., & Hayne, H. (1993). The time window hypothesis: Implications for categorization and memory modification. *Infant Behavior and Development*, 16, 139-157.
- Rovee-Collier, C., Evancio, S., & Earley, L. A. (1995). The time window hypothesis: Spacing effects. *Infant Behavior and Development*, 18, 69-78.
- Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, 26, 709-777.
- Schapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal*, 41, 159-189.

- Scharf, M. T., Woo, N. H., Lattal, K. M., Nguyen, P. V., & Abel, T. (2002). Protein synthesis is required for the enhancement of long-term potentiation and long-term memory by spaced training. *Journal of Neurophysiology*, 87, 2770-2777.
- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York, NY: Basic Books.
- Skinner, B. F. (1959). *Cumulative record*. New York, USA: Appleton-Century-Crofts.
- Slavin, R. E. (2002). Evidence based education policies: Transforming educational practice and research. *Educational Researcher*, 31, 15-21.
- Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, 25, 763-767. doi: 10.1002/acp.1747
- Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year olds. *Journal of Experimental Child Psychology*, 109, 132-140.
- Sulzer-Azaroff, B. & Mayer, G. R. (1991). *Behavior analysis for lasting change*. Fort Worth, TX: Holt, Rhinehart & Winston.
- Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, 3. doi: 10.3389/psyq.2012.0046.
- Waxman, S. R., & Booth, A. E. (2000). Principles that are invoked in the acquisition of words but not facts. *Cognition*, 77, 33-43.
- Zhou, Q., Tao, H. W., & Poo, M. M. (2003). Reversal and stabilization of synaptic modifications in a developing visual system. *Science*, 300, 1953-1957.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3-17.
- Zull, J.E. (2011). *From brain to mind: Using neuroscience to guide change in education*. Sterling, VA: Stylus Publishing.

APPENDIX 1

University of Canterbury Human Ethics Approval Letters



Human Ethics Committee

Tel: +64 3 364 2241, Fax: +64 3 364 2856, Email: human-ethics@canterbury.ac.nz

Ref: 2010/08/ERHEC

9 April 2010

Brett Clark  
School of Educational Studies & Human Development  
College of Education  
UNIVERSITY OF CANTERBURY

Dear Brett

Thank you for providing the revised documents in support of your application to the Educational Research Human Ethics Committee. I am very pleased to inform you that your research proposal "A study of the effects on retention of different time intervals between opportunities to learn" has been granted ethical approval.

Please note that should circumstances relevant to this current application change you are required to reapply for ethical approval.

If you have any questions regarding this approval please let me know.

We wish you well for your research.

Yours sincerely

  
Dr Missy Morton  
Chair  
Educational Research HEC

*"Please note that Ethical Approval relates only to the ethical elements of the relationship between the researcher, research participants and other stakeholders. The granting of approval or clearance by the Educational Research Human Ethics Committee should not be interpreted as comment on the methodology, legality, value or any other matters relating to this research."*

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand [www.canterbury.ac.nz](http://www.canterbury.ac.nz)



Human Ethics Committee

Tel: 064 3 364 2247, Fax: 064 3 364 2858, Email: human.ethics@canterbury.ac.nz



Ref: HEC 2010/08/ERHEC

8 August 2011

Brett Clark  
St Andrews College  
347 Papanui Road  
CHRISTCHURCH

Dear Brett,

I am very pleased to inform you that your application for an amendment to the research proposal "A study of the effects on retention of different time intervals between opportunities to learn" has been granted ethical approval from the Educational Research Human Ethics Committee.

Please note that should circumstances relevant to this current application change you are required to reapply for ethical approval.

If you have any questions regarding this approval please let me know.

We wish you well for your continuing research.

Yours sincerely

A handwritten signature in blue ink, appearing to read 'Nicola Surtees', with the initials 'PP' written to the left.

Nicola Surtees  
**Chair**  
**Educational Research Human Ethics Committee**

*"Please note that Ethical Approval and/or Clearance relates only to the ethical elements of the relationship between the researcher, research participants and other stakeholders. The granting of approval or clearance by the Educational Research Human Ethics Committee should not be interpreted as comment on the methodology, legality, value or any other matters relating to this research."*

University of Canterbury Press: Bag 4800, Christchurch 8140, New Zealand [www.canterbury.ac.nz](http://www.canterbury.ac.nz)

Human Ethics Committee

Tel: +64 3 364 2241, Fax: +64 3 364 2856, Email: [human.ethics@canterbury.ac.nz](mailto:human.ethics@canterbury.ac.nz)



Ref: 2010/08/ERHEC

2 February 2012

Brett Clark  
School of Educational Studies & Human Development  
College of Education  
UNIVERSITY OF CANTERBURY

Dear Brett

I am very pleased to inform you that your application for an amendment to the research proposal "A study of the effects on retention of different time intervals between opportunities to learn" has been granted ethical approval from the Educational Research Human Ethics Committee.

Please note that should circumstances relevant to this current application change you are required to reapply for ethical approval.

If you have any questions regarding this approval please let me know.

We wish you well for your continuing research.

Yours sincerely

PP 

Nicola Sutcliffe  
Chair  
Educational Research Human Ethics Committee

*"Please note that Ethical Approval and/or Clearance relates only to the ethical elements of the relationship between the researcher, research participants and other stakeholders. The granting of approval or clearance by the Educational Research Human Ethics Committee should not be interpreted as comment on the methodology, legality, value or any other matters relating to this research."*

## APPENDIX 2

### Student Information Sheet and Student Consent Form



Tel: +64 3 , Fax: +64 3  
Email: @canterbury.ac.nz

#### Student information sheet

##### The effects on retention of different time intervals between opportunities to learn

The aim of the study you have been taking part in was to test a theory of learning developed by Professor Graham Nuthall and Adrienne Allon-Lee which states that for a student to learn a new concept the student must work with that concept on at least four occasions with no more than a two day gap between any two of those learning experiences. For my PHD I am exploring the kinds of content for which this finding applies and the circumstances under which it is effective. Knowing when and where this rule predicts student learning has major implications for teachers and learners.

As is customary you sat a pre-test for the Maths unit to determine what you did and did not know about the material to be learned for that unit. I selected half a dozen or so concepts and facts from the pre-test that you did not know and then tracked your learning of those concepts and facts through to the post-test.

Obviously I could not tell you in advance which concepts and facts I was tracking during the experimental lessons. For the experiment to work I needed to keep the fact that I was conducting an experiment and the content of the experiment secret.

Yours sincerely,

Brett Clark  
40 Balrudry St.  
Avonhead  
Christchurch

Dr. John Church  
School of Educational Studies and Human Development  
University of Canterbury

Tel: +64 3 , Fax: +64 3  
Email: [uc@canterbury.ac.nz](mailto:uc@canterbury.ac.nz)



### Student debriefing sheet

#### The effects on retention of different time intervals between opportunities to learn

Dear friend,

You will be interested to know that I have been studying your learning during the Maths unit by looking to see if your knowledge of particular concepts has improved from the pre-test to the post-test. I am required by the university to tell you that I am planning to use your test scores on the Maths unit in the thesis that I am writing for the university as part of my PHD. Of course the thesis will refer to participants only by their code number. It will not refer to people's names or the name of the school or any other identifying information. To ensure confidentiality and your anonymity the following steps have been taken:

- i. The data sheets only carry a code number for the student and not your name.
- ii. The number coded student data sheets are being stored in a locked filing cabinet.
- iii. The data gathered will be destroyed after five years.

Please note that the information could be used for publication. I will also provide a summary of the results to all participants.

Attached is a consent form in which you indicate whether or not you are willing to participate in the research project.

---

### Student consent form

Please indicate whether you are willing to allow Mr. Clark to use the data he has gathered for his research by ticking the appropriate box.

☐ YES

☐ NO

After completing the form hand the form to me in class at the end of the period.

## APPENDIX 3

### Parent Information Sheet and Parent Consent Form

Tel: +64 3 , Fax: +64 3  
Email: @canterbury.ac.nz



### Parental Consent form

#### The effects on retention of different time intervals between opportunities to learn

##### Declaration of Consent to Participate

I have read and understood the information provided about this research project.

I understand that participation in the experiment will not place additional academic demands on participants.

I understand that my son/daughter might be involved in more than one experiment.

I understand that my son/daughter will be debriefed after the experiment and at this time he/she has the right to refuse his/her data from being used for the research.

I understand that the use of code numbers for students and not students' names will ensure the anonymity of the participants.

I understand that there will be only a single copy of the key which matches the code numbers to student names and that Mr. Clark will keep this safe in my pass-worded computer.

I understand that all data sheets will be securely stored in a locked filing cabinet and that all data will be destroyed after five years.

I understand that the experiments will take place over the course of the academic year.

I understand that the data gathered will be destroyed after five years.

I understand that Mr. Clark will provide all participants with a summary of results.

Finally, I understand that the information could be used for publication.

Please indicate below whether you will consent to your child participating in the research.

### Parental consent form

Please indicate whether or not you wish your child's data to be used for the research project by ticking either the "Yes" or "No" box below.

☐

Yes

☐

No

Parent's signature:

Date:

## Parents' Information sheet

### The effects on retention of different time intervals between opportunities to learn

In the last few weeks at the College your child, as part of their normal lessons in [subject], has been participating in my PHD research study. The aim of the study is to test a theory of learning developed by Professor Graham Nuthall and Adrienne Alton-Lee which states that for a student to learn a new concept the student must work with that concept on at least four occasions with no more than a two day gap between any two of those learning experiences. For my PHD I am exploring the kinds of content for which this finding applies and the circumstances under which it is effective. Knowing when and where this rule predicts student learning has major implications for teachers and learners.

As is customary your child sat a pre-test for the XXX unit to determine what he/she did and did not know about the material to be learned for that unit. I selected half a dozen or so concepts and facts from the pre-test that he/she did not know and then tracked his/her learning of those concepts and facts through to the post-test.

Obviously I could not tell your son/daughter in advance which concepts and facts I was tracking during the experimental lessons. For the experiment to work I needed to keep the fact that I was conducting an experiment and the content of the experiment secret.

Yours sincerely,

Brett Clark  
40 Balrudry St  
Avonhead  
Christchurch

Dr. John Church  
School of Educational Studies and Human Development  
University of Canterbury

## APPENDIX 4

### Test Scores for Each Participant for Each Experimental Treatment

#### *Test Scores of Participants in Treatment 1 of Experiment 1*

Name of participant	Pre-test score	Post-test score	Real score	% score
Sam	0/14	13/14	13/14	93%
Ben	0/14	9/14	9/14	64%
Fred	0/14	13/14	13/14	93%
Mark	0/14	14/14	14/14	100%
Sarah	0/14	14/14	14/14	100%
Seb	2/14	14/14	12/12	100%
Vicky	0/14	13/14	13/14	93%
Clark	0/14	13/14	13/14	93%
Andy	4/14	14/14	10/10	100%
Olivia	1/14	14/14	13/13	100%
Libby	1/14	14/14	13/13	100%
Sally-anne	1/14	13/14	12/13	92%
Angus	2/14	14/14	12/12	100%
Harry	0/14	13/14	13/14	93%
Susan	1/14	14/14	13/13	100%
Lewis	1/14	14/14	13/13	100%
Heidi	0/14	13/14	13/14	93%
Ellie	0/14	12/14	12/14	86%



*Test Scores of Participants in Treatment 2 of Experiment 1*

Name of participant	Pre-test score	Post-test score	Real score	% score
Seb	0/14	12/14	12/14	86%
Fred	1/14	13/14	12/13	92%
Ben	0/14	13/14	13/14	93%
Sally-Anne	0/14	10/14	10/14	71%
Libby	0/14	13/14	13/14	93%
Mark	0/14	14/14	14/14	100%
Ellie	0/14	9/14	9/14	64%
Vicky	0/14	13/14	13/14	93%
Andy	5/14	14/14	9/9	100%
Susan	0/14	14/14	14/14	100%
Olivia	0/14	14/14	14/14	100%
Heidi	0/14	14/14	14/14	100%
Harry	2/14	14/14	12/12	100%
Sam	4/14	14/14	10/10	100%
Lewis	0/14	13/14	13/14	93%
Sarah	0/14	8/14	8/14	57%
Clark	1/14	10/14	9/13	69%
Angus	2/14	14/14	12/12	100%
Ian	2/14	14/14	12/12	100%

*Test Scores of Participants in Treatment 3 of Experiment 1*

Name of participant	Pre-test score	Post-test score	Real score	% score
Harry	0/14	10/14	10/14	71%
Ben	1/14	7/14	6/13	46%
Lewis	0/14	8/14	8/14	57%
Olivia	0/14	13/14	13/14	93%
Mark	0/14	13/14	13/14	93%
Seb	0/14	10/14	10/14	71%
Fred	0/14	8/14	8/14	57%
Vicky	0/14	10/14	10/14	71%
Susan	0/14	13/14	13/14	93%
Libby	0/14	11/14	11/14	79%
Sam	0/14	9/14	9/14	64%
Angus	0/14	12/14	12/14	86%
Sally-Anne	0/14	6/14	6/14	43%
Andy	1/14	11/14	10/13	77%
Sarah	0/14	8/14	8/14	57%
Heidi	0/14	6/14	6/14	43%
Clark	0/14	6/14	6/14	43%
Ian	0/14	8/14	8/14	57%

*Test Scores of Participants in Treatment 1 of Experiment 2*

Name of participant	Pretest score	Post-test score	Real score	% score
Livi	1/12	12/12	11/11	100%
Emily	1/12	12/12	11/11	100%
Peter	2/12	11/12	9/10	90%
Rachel	1/12	11/12	10/11	91%
Liam	1/12	11/12	10/11	91%
Karen	1/12	12/12	11/11	100%
Reana	1/12	10/12	9/11	82%
Charles	2/12	11/12	9/10	90%
Maddy	0/12	11/12	11/12	92%
Grace	2/12	12/12	10/10	100%
Paulina	2/12	11/12	9/10	90%
Bianca	1/12	12/12	11/11	100%
Lucy	2/12	12/12	10/10	100%
Molly	2/12	12/12	10/10	100%
Will	2/12	12/12	10/10	100%
James	1/12	10/12	9/11	82%
Andrew	1/12	10/12	9/11	82%
Sandra	0/12	11/12	11/12	92%
Monica	2/12	12/12	10/10	100%
Jane	1/12	12/12	11/11	100%

*Test Scores of Participants in Treatment 2 of Experiment 2*

Name of participant	Pretest score	Post-test score	Real score	% score
Jane	1/11	9/11	8/10	80
Livi	1/11	11/11	10/10	100%
Emily	2/11	11/11	9/9	100%
Peter	1/11	11/11	10/10	100%
Sandra	2/11	11/11	9/9	100%
Liam	1/11	11/11	10/10	100%
Karen	0/11	11/11	11/11	100%
Reana	0/11	7/11	7/11	64%
Andrew	3/11	11/11	8/8	100%
Molly	0/11	11/11	11/11	100%
Charles	1/11	10/11	9/10	90%
Maddy	0/11	10/11	10/11	91%
Lucy	2/11	11/11	9/9	100%
Paulina	0/11	11/11	11/11	100%
Will	1/11	11/11	10/10	100%

*Test Scores of Participants in Treatment 3 of Experiment 2*

Name of participant	Pretest score	Post-test score	Real Score	% score
Jane	0/12	5/12	5/12	42%
Livi	0/12	9/12	9/12	75%
Emily	1/12	7/12	6/11	55%
Peter	0/12	9/12	9/12	75%
Rachel	0/12	9/12	9/12	75%
Liam	0/12	11/12	11/12	92%
Karen	0/12	8/12	8/12	67%
Reana	0/12	3/12	3/12	25%
Andrew	1/12	8/12	7/11	64%
Charles	0/12	8/12	8/12	67%
Maddy	0/12	5/12	5/12	42%
Grace	1/12	5/12	4/11	36%
Sandra	0/12	10/12	10/12	83%
Bianca	0/12	7/12	7/12	58%
Lucy	1/12	6/12	5/11	45%
Monica	1/12	8/12	7/11	64%
Molly	2/12	9/12	7/10	70%
Will	0/12	11/12	11/12	92%
James	0/12	8/12	8/12	67%

*Test Scores of Participants in Treatment 1 of Experiment 3*

Name of participant	Pretest score	Post-test score	Real score	% score
Thomas	0/14	13/14	13/14	93%
Sumant	0/14	12/14	12/14	86%
Kerry	0/14	13/14	13/14	93%
Stephanie	0/14	13/14	13/14	93%
Ross	0/14	13/14	13/14	93%
Hugh	3/14	14/14	11/11	100%
Penny	8/14	13/14	5/6	83%

*Test Scores of Participants in Treatment 2 of Experiment 3*

Name of participant	Pretest score	Post-test score	Real score	% score
Thomas	0/12	11/12	11/12	92%
Sumant	0/12	12/12	12/12	100%
Hugh	0/12	10/12	10/12	83%
Kerry	0/12	12/12	12/12	100%
Penny	0/12	12/12	12/12	100%
John	0/12	12/12	12/12	100%
Ross	0/12	12/12	12/12	100%
Rob	1/12	12/12	11/11	100%
Stephanie	0/12	10/12	10/12	83%

*Test Scores for Participants in Treatment 3 of Experiment 3*

Name of participant	Pretest score	Post-test score	Real score	% score
Ross	0/13	13/13	13/13	100%
Stephanie	0/13	13/13	13/13	100%
Kerry	5/13	12/13	7/8	88%
Sumant	0/13	5/13	5/13	38%
John	7/13	13/13	6/6	100%
Rob	4/13	13/13	9/9	100%
Thomas	0/13	12/13	12/13	92%



*Test Scores of Participants in Treatment 4 of Experiment 3*

Name of participant	Pretest score	Post-test score	Real score	% score
Ross	0/13	13/13	13/13	100%
Stephanie	0/13	13/13	13/13	100%
Hugh	0/13	13/13	13/13	100%
Rob	0/13	13/13	13/13	100%
John	0/13	12/13	12/13	92%
Kerry	0/13	12/13	12/13	92%
Penny	0/13	10/13	10/13	77%
Thomas	0/13	9/13	9/13	69%

*Test Scores of Participants in Treatment 1 of Pilot Study 1*

Name of participant	Pretest score	Post-test score	Real score	% score
Tamsin	0/13	13/13	13/13	100%
Sarah	0/13	13/13	13/13	100%
Kelly	0/13	13/13	13/13	100%
Martina	0/13	13/13	13/13	100%
Finn	0/13	13/13	13/13	100%
Claire	0/13	12/13	12/13	92%
James	0/13	13/13	13/13	100%
Penelope	0/13	13/13	13/13	100%
Didi	0/13	13/13	13/13	100%
Roger	0/13	13/13	13/13	100%
Lawrence	0/13	13/13	13/13	100%
Joel	0/13	13/13	13/13	100%

*Test Scores of Participants in Treatment 2 of Pilot Study 1*

Name of participant	Pretest score	Post-test score	Real score	& score
Tamsin	0/13	4/13	4/13	31%
Sarah	0/13	10/13	10/13	77%
Kelly	2/13	5/13	3/11	27%
Martina	3/13	5/13	2/10	20%
Finn	0/13	3/13	3/13	23%
Claire	0/13	5/13	5/13	38%
James	3/13	5/13	2/10	20%
Penelope	0/13	5/13	5/13	38%
Didi	0/13	5/13	5/13	38%
Roger	0/13	3/13	3/13	23%
Tamara	0/13	11/13	11/13	85%
Lawrence	0/13	4/13	4/13	31%
Joel	0/13	1/13	1/13	8%

*Test Scores for Treatment 3 of Pilot Study 1*

Name of participant	Pretest score	Post-test score	Real score	5 score
Tamsin	0/13	7/13	7/13	54%
Joel	0/13	1/13	1/13	8%
Kelly	1/13	2/13	1/12	8%
Martina	2/13	5/13	3/11	27%
Finn	1/13	5/13	4/12	33%
Claire	2/13	9/13	7/11	64%
James	1/13	8/13	7/12	58%
Penelope	1/13	13/13	12/12	100%
Didi	1/13	13/13	12/12	100%
Roger	0/13	6/13	6/13	46%
Tamara	0/13	5/13	5/13	38%

*Test Scores of Participants in Treatment 1 of Experiment 4*

Name of participant	Pretest score	Post-test score	Real score	% score
Jane	0/12	11/12	11/12	92%
Livi	2/12	12/12	10/10	100%
Emily	2/12	11/12	9/10	90%
Peter	1/12	11/12	10/11	91%
Rachel	0/12	12/12	12/12	100%
Liam	3/12	12/12	9/9	100%
Reana	0/12	9/12	9/12	75%
Linda	1/12	10/12	9/11	82%
Paulina	1/12	11/12	10/11	91%
Lucy	1/12	11/12	10/11	91%
Will	0/12	11/12	11/12	92%

*Test Scores of Participants in Treatment 2 of Experiment 4*

Name of participant	Pretest score	Post-test score	Real score	% score
Jane	0/12	11/12	11/12	92%
Livi	0/12	12/12	12/12	100%
Emily	0/12	11/12	11/12	92%
Peter	0/12	11/12	11/12	92%
Rachel	0/12	12/12	12/12	100%
Liam	0/12	12/12	12/12	100%
Reana	0/12	10/12	10/12	83%
Linda	0/12	11/12	11/12	92%
Paulina	1/12	12/12	11/11	100%
Lucy	0/12	12/12	12/12	100%
Will	0/12	12/12	12/12	100%

*Test Scores of Participants in Treatment 1 of Experiment 5*

Name of participant	Pretest score	Post-test score	Real score	% score
Cindy	4/12	11/12	7/8	88%
Maggie	0/12	11/12	11/12	92%
Dillon	2/12	12/12	10/10	100%
Ben	1/12	11/12	10/11	91%
Eliza	5/12	12/12	7/7	100%
Harriet	2/12	11/12	9/10	90%
Joanne	0/12	11/12	11/12	92%
Georgia	1/12	12/12	11/11	100%
Bond	0/12	10/12	10/12	83%
Alison	5/12	11/12	6/7	86%
Genevieve	2/12	11/12	9/10	90%
Simone	3/12	9/12	6/9	67%

*Test Scores of Participants in Treatment 2 of Experiment 5*

Name of participant	Pretest score	Post-test score	Real score	% score
Joanne	1/12	8/12	7/11	64%
Mel	1/12	11/12	10/11	91%
Dillon	2/12	11/12	9/10	90%
Georgia	3/12	11/12	8/9	89%
Alison	7/12	12/12	5/5	100%
Harriet	2/12	12/12	10/10	100%
Simone	2/12	9/12	7/10	70%
Eliza	2/12	11/12	9/10	90%
Tania	2/12	11/12	9/10	90%
Beatrice	2/12	11/12	9/10	90%
Genevieve	2/12	12/12	10/10	100%
Maggie	1/12	8/12	7/11	64%
Cindy	5/12	11/12	6/7	86%
Bond	1/12	11/12	10/11	91%



*Test Scores of Participants in Treatment 3 of Experiment 5*

Name of participant	Pretest score	Post-test score	Real score	& score
Maggie	0/12	12/12	12/12	100%
Harriet	2/12	11/12	9/10	90%
Bond	1/12	9/12	8/11	73%
Alison	7/12	12/12	5/5	100%
Tania	3/12	12/12	9/9	100%
Ben	0/12	9/12	9/12	75%
Simone	1/12	10/12	9/11	82%
Genevieve	1/12	11/12	10/11	91%
Beatrice	1/12	9/12	8/11	73%
Cindy	1/12	12/12	11/11	100%
Eliza	2/12	12/12	10/10	100%
Georgia	1/12	11/12	10/11	91%
Mel	0/12	7/12	7/12	58%
Dillon	2/12	10/12	8/10	80%
Joanne	0/12	9/12	9/12	75%

*Test Scores of Participants in Treatment 1 of Experiment 6*

Name of participant	Pretest score	Post-test score	Real score	% score
William	2/13	13/13	11/11	100%
Joe	0/13	13/13	13/13	100%
Cathy	0/13	12/13	12/13	92%
Margaret	0/13	12/13	12/13	92%
Cam	0/13	13/13	13/13	100%
Eugene	0/13	13/13	13/13	100%
Debbie	0/13	12/13	12/13	92%
Hayley	1/13	12/13	11/12	92%
Angela	1/13	13/13	12/12	100%
Bindy	0/13	11/13	11/13	85%
Sally	0/13	12/13	12/13	92%
Jeremy	2/13	13/13	11/11	100%
Juliana	2/13	12/13	10/11	91%
Ken	1/13	13/13	12/12	100%

*Test Scores of Participants in Treatment 2 of Experiment 6*

Name of participant	Pretest score	Post-test score	Real score	% score
Joe	0/14	13/14	13/14	93%
Margaret	0/14	14/14	14/14	100%
Jeremy	1/14	12/14	11/13	85%
Juliana	0/14	11/14	11/14	79%
Bindy	0/14	10/14	10/14	71%
Eugene	0/14	11/14	11/14	79%
Angela	0/14	14/14	14/14	100%
Debbie	0/14	14/14	14/14	100%
William	1/14	14/14	13/13	100%
Hayley	2/14	14/14	12/12	100%
Cam	0/14	7/14	7/14	50%
Ken	0/14	11/14	11/14	79%
Sally	0/14	12/14	12/14	86%
Cathy	1/14	14/14	13/13	100%

*Test Scores of Participants in Treatment 1 of Experiment 7*

Name of participant	Pretest score	Post-test score	Real score	% score
Belinda	2/13	13/13	11/11	100%
Andrea	2/13	13/13	11/11	100%
Madison	6/13	13/13	7/7	100%
Samantha	8/13	13/13	5/5	100%
Neil	1/13	11/13	10/12	83%
Luke	1/13	11/13	10/12	83%
Hector	2/13	11/13	9/11	82%
Fergus	4/13	13/13	9/9	100%
Johanna	5/13	12/13	7/8	88%
Sue	4/13	13/13	9/9	100%
Steve	2/13	13/13	11/11	100%
Jim	2/13	13/13	11/11	100%
Emma	5/13	12/13	7/8	88%

*Test Scores of Participants in Treatment 2 of Experiment 7*

Name of participant	Pretest score	Post-test score	Real score	% score
Steve	4/12	12/12	8/8	100%
Fergus	4/12	12/12	8/8	100%
Hector	0/12	11/12	11/12	92%
Samantha	5/12	12/12	7/7	100%
Luke	3/12	10/12	7/9	78%
Emma	2/12	11/12	9/10	90%
Andrea	0/12	8/12	8/12	67%
Jim	3/12	10/12	7/9	78%
Neil	2/12	10/12	8/10	80%
Sue	4/12	12/12	8/8	100%
Belinda	0/12	9/12	9/12	75%
Johanna	4/12	11/12	7/8	88%
Madison	5/12	12/12	7/7	100%

*Test Scores for Participants in Treatment 1 of Experiment 8*

Name of participant	Pretest score	Post-test score	Real score	% score
Kirsty	3/12	12/12	9/9	100%
Katrina	0/12	10/12	10/12	83%
Jerome	0/12	11/12	11/12	92%
Mary	1/12	12/12	11/11	100%
Nicola	2/12	11/12	9/10	90%
Abby	1/12	9/12	8/11	73%
Jack	2/12	11/12	9/10	90%
Max	0/12	10/12	10/12	83%
Henrietta	2/12	11/12	9/10	90%
Lucia	3/12	11/12	8/9	89%
Anastasia	1/12	12/12	11/11	100%
Kim	1/12	11/12	10/11	91%
Anna	1/12	12/12	11/11	100%
Jessica	2/12	11/12	9/10	90%
Kate	1/12	11/12	10/11	91%
Christine	2/12	11/12	9/10	90%

*Test Scores of Participants in Treatment 2 of Experiment 8*

Name of participant	Pretest score	Post-test score	Real score	% score
Nicola	0/16	14/16	14/16	88%
Anastasia	0/16	13/16	13/16	81%
Kirsty	3/16	16/16	13/13	100%
Jessica	0/16	11/16	11/16	69%
Anna	0/16	15/16	15/16	94%
Kim	0/16	15/16	15/16	94%
Mary	0/16	15/16	15/16	94%
Henrietta	0/16	15/16	15/16	94%
Jack	0/16	14/16	14/16	88%
Katrina	0/16	15/16	15/16	94%
Jerome	0/16	11/16	11/16	69%
Lucia	0/16	14/16	14/16	88%
Christine	0/15	15/16	15/16	94%
Kate	0/16	14/16	14/16	88%
Max	0/16	15/16	15/16	94%
Abby	0/16	6/16	6/16	38%

*Test Scores of Participants in Treatment 1 of Experiment 9*

Name of participant	Pretest score	Post-test score	Real score	% score
Perry	2/13	12/13	10/11	92%
Phillipa	2/13	13/13	11/11	100%
Mercy	0/13	12/13	12/13	92%
Maryanne	0/13	11/13	11/13	85%
Bevan	2/13	13/13	11/11	100%
Dan	2/13	13/13	11/11	100%
Brad	0/13	13/13	13/13	100%
Oscar	0/13	12/13	12/13	92%
Zach	0/13	11/13	11/13	85%
Corey	0/13	10/13	10/13	77%
Tane	0/13	13/13	13/13	100%
Simon	2/13	12/13	10/11	91%



*Test Scores of Participants in Treatment 2 of Experiment 9*

Name of participant	Pretest score	Post-test score	Real score	% score
Simon	0/14	14/14	14/14	100%
Maryanne	2/14	10/14	8/12	67%
Dan	0/14	10/14	10/14	71%
Brad	0/14	13/14	13/14	93%
Phillipa	0/14	7/14	7/14	50%
Zach	0/14	11/14	11/14	79%
Mercy	0/14	14/14	14/14	100%
Corey	0/14	8/14	8/14	57%
Tane	0/14	11/14	11/14	79%
Perry	1/14	12/14	11/13	85%
Bevan	1/14	14/14	13/13	100%
Oscar	0/14	6/14	6/14	43%

*Test Scores for Participants in Treatment 1 of Experiment 10*

Name of participant	Pretest score	Post-test score	Real score	% score
Ivan	4/14	13/14	9/10	90%
Angus	5/14	13/14	8/9	89%
Jordan	0/14	13/14	13/14	93%
David	1/14	13/14	12/13	92%
Melanie	1/14	13/14	12/13	92%
Ernest	1/14	13/14	12/13	92%
Gerald	4/14	14/14	10/10	100%
Tom	0/14	14/14	14/14	100%
Matt	1/14	14/14	13/13	100%
Louis	1/14	14/14	13/13	100%
Derek	5/14	14/14	9/9	100%
Aimee	0/14	13/14	13/14	93%

*Test Scores for Participants in Treatment 2 of Experiment 10*

Name of participant	Pretest score	Post-test score	Real score	% score
Ivan	0/12	12/12	12/12	100%
Angus	0/12	12/12	12/12	100%
Jordan	0/12	11/12	11/12	92%
David	0/12	12/12	12/12	100%
Melanie	0/12	12/12	12/12	100%
Ernest	1/12	11/12	10/11	91%
Gerald	0/12	12/12	12/12	100%
Tom	0/12	11/12	11/12	92%
Matt	1/12	12/12	11/11	100%
Louis	0/12	12/12	12/12	100%
Derek	0/12	12/12	12/12	100%
Aimee	1/12	12/12	11/11	100%
Huw	0/12	9/12	9/12	75%

*Test Scores for Participants in Treatment 3 of Experiment 10*

Name of participant	Pretest score	Post-test score	Real score	% score
Ivan	1/13	13/13	12/12	100%
Melanie	0/13	12/13	12/13	92%
Gerald	7/13	13/13	6/6	100%
Tom	8/13	11/13	3/5	60%
Matt	7/13	13/13	6/6	100%
Derek	7/13	13/13	6/6	100%
Aimee	7/13	13/13	6/6	100%
Huw	6/13	13/13	13/13	100%

*Note.* Ivan and Melanie had an 11-day/7-day/2-day schedule of ISIs. The other participants had an 11-day/3-day/6-day schedule of ISIs.