Reframing the Transition:

A Mixed Methods Case Study of Mathematical Thinking Workshops in First-Year University Mathematics

A Doctoral Thesis

Department of Mathematics and Applied Mathematics

Faculty of Science

University of Cape Town

Submitted in September 2025

Presented By

Mashudu Mokhithi

MSc, BSc (Eng)

for the degree of

Doctor of Philosophy

Supervised by:

Associate Professor Anita L. Campbell

Associate Professor Jonathan P. Shock

Associate Professor Pragashni Padayachee

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

DECLARATIONS

Plagiarism

I, Mashudu Mokhithi, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorise the University to reproduce for the purpose of research either the whole or any portion of the contents in any matter whatsoever.

Use of generative AI

I acknowledge the use of OpenAI's ChatGPT (GPT-4, April 2024) strictly for improving readability and style and to correct grammar, spelling, punctuation, and tone. Its use was limited to wording and formatting suggestions and did not involve autonomous content creation. All ideas, analyses, and conclusions are my own; I reviewed and approved every edit and retain full responsibility for the final text. No generative AI tools were used to generate original research data.

Signed by candidate

10/12/2025

Candidate : Mashudu Mokhithi Date

ACKNOWLEDGEMENTS

Above all, I give thanks to Christ, for it is in His strength that I found the courage and peace to complete this work, as Apostle Paul reminds us in Philippians 4:13.

I am deeply grateful to my supervisors, without whom this work would not have been possible. Associate Prof. Anita Campbell, thank you for guiding me through the writing process, believing in me when things were tough, and pushing me to do better. Associate Prof. Jonathan Shock, thank you for guiding the quantitative data analysis and for your constructive and rigorous critique. Associate Prof. Pragashni Padayachee, thank you for helping me bring the theoretical strands together, write more cohesively, and anticipate what examiners would look for. I am grateful to my supervisors for far more than I can write here.

I also thank the Mathematical Thinking Workshops design team, especially Associate Prof. Jonathan Shock, Associate Prof. Kate le Roux, Neil Eddy, and Lizelle Nitt, as well as those who were involved in the early stages, including Dr Harry Wiggins, Associate Prof. Bob Osano, and Dr Claire Blackman. Many thanks also go to the workshop participants who volunteered for focus group interviews; you provided rich insights that I hope will have lasting impacts. I am indebted to Nazeema Ahmed, who conducted all the focus group interviews voluntarily.

Finally, I would like to extend my deepest thanks to my family and friends for their unwavering support and words of encouragement.

ABSTRACT

For many students, the transition from school to university mathematics is challenging, especially in systems with persistent inequality, such as South Africa. In South Africa, the transition is further shaped by the enduring effects of apartheid on schooling opportunities and preparation for many learners. Support initiatives exist, but most emphasise procedural skills and content mastery while underemphasising skills essential for university mathematics, such as advanced mathematical thinking and metacognitive strategies.

This study examines a theoretically grounded alternative that prioritises higher-order thinking and self-regulation. Specifically, it evaluates a voluntary, supplemental set of Mathematical Thinking Workshops (MTWs), which were offered each semester and consisted of two two-hour sessions per week. The workshops foreground collaborative problem-solving, structured reflection, and the unpacking of conceptual structures, and are grounded in the Zone of Proximal Development (ZPD), Cognitive Load Theory (CLT), and APOS theory, with the aim of cultivating the mathematical thinking and metacognitive abilities of first-year BSc students.

A convergent parallel mixed-methods design integrated quantitative and qualitative strands. Quantitatively, propensity score methods were employed to mitigate confounding due to differences in pre-university academic achievement between the treatment (workshop participants) and control groups. The Average Treatment Effect on the Treated (ATT) was then estimated with cluster-robust standard errors and Holm-adjusted p-values. Analyses also examined higher-order performance (HO%) and heterogeneity by NSC Mathematics bands (70–84 vs 85-100), with Rosenbaum bounds used as a sensitivity check for unobserved confounding. Qualitatively, a reflexive thematic analysis of six focus group discussions explored the perceived impacts on mathematical understanding, problem-solving confidence, and learning strategies.

Quantitatively, participants outperformed matched non-participants on early first-semester assessments and higher-order tasks, with the effects of the final exams being

more minor and more variable under cluster-robust, Holm-adjusted inference. Qualitatively, participants described shifts from procedural to conceptual reasoning, strengthened metacognitive regulation (planning, monitoring, and reflection), increased confidence, and a greater sense of belonging within the mathematical community. Facilitator practices (supportive, low-threat climate) and peer collaboration emerged as key mechanisms for success. Collectively, the findings provide a theoretically informed, evidence-based framework for mathematics support that targets higher-order cognition and metacognition, and identify a subgroup (i.e., students in the 70–84 NSC band) for whom MTWs may be especially beneficial.

DEFINITION OF TERMS

APOS Theory: A theoretical model of cognitive development in mathematics, which proposes that understanding evolves through four mental structures: Action (algorithmic manipulation of mathematical objects), Process (internalisation of actions to form coherent mental procedures), Object (conceptualisation of processes as abstract entities that can be manipulated), and Schema (integration of multiple objects and processes into a flexible and generalised mental framework) (Asiala et al., 1996).

Average Treatment Effect for the Overlap Population (ATO): The ATO represents the causal effect of a treatment on the overlap population, i.e., students whose covariates make both the treatment and control plausible. The ATO addresses challenges in causal inference when treatment and control groups have limited overlap in covariates. (Orihara et al., 2024).

Average Treatment Effect on the Treated (ATT): An estimate of the mean difference in outcomes between those who participated in the intervention (workshops) and what their outcomes would have been had they not participated. ATT is used to assess the impact of the workshops while accounting for non-random assignment (Shu & Tan, 2018).

Caliper Matching: A restriction in nearest-neighbour matching (see Propensity Score Matching definition) that only allows matches within a maximum distance ("caliper"). In this study, 0.2 of the Standard Deviation on the propensity score. Pairs exceeding the caliper were discarded. (Kormos et al., 2023).

Cluster-Robust Standard Errors: Standard errors adjusted for dependence among observations with clusters (e.g., when the same control is reused across multiple matches).

Cognitive Load Theory (CLT): A theory of instructional design that highlights the limitations of working memory. It distinguishes between intrinsic, extraneous, and

germane cognitive load and informs how workshop tasks are sequenced to maximise conceptual learning without overloading students (Sweller, 1988).

Common Support/Overlap: There is a substantial overlap in the distribution of propensity scores between the treatment group (those who attended the workshops) and the control group (those who did not attend the workshops).

Constructivism: A learning theory asserting that knowledge is actively constructed by learners through experience, reflection, and interaction, rather than passively received (von Glasersfeld, 1995). Constructivist principles form the basis of the student-centred, problem-based structure of the workshops.

Covariates: Covariates are pre-existing variables that may influence both treatment assignment and outcomes in a study. In the context of this research, covariates refer to students' background characteristics (National Senior Certificate [NSC] scores and National Benchmark Test [NBT] scores) that are used to control for selection bias when estimating the effect of workshop participation.

Familywise Error Rate (FWER): The probability of making at least one Type I error across a family of related tests. Controlled here using Holm's procedure.

Holm adjustment: The Holm sequential Bonferroni procedure is a less conservative alternative to the traditional Bonferroni method for controlling familywise error rate in multiple testing scenarios. (Eichstaedt et al., 2013).

Learner/student: In this thesis, *student* will refer to an individual formally enrolled in a course or programme at a university or school, whereas *learner* will be used more generally for any person engaged in learning (including students). Accordingly, occasional interchange will occur when the context shifts from describing the enrolled *students* to discussing learning processes or theories. In South Africa, however, *learners* commonly denotes school pupils, especially at high-school level.

Propensity Score: A propensity score, estimated using logistic regression in this study, is the probability that a unit (e.g., a student) receives a treatment or intervention, given a set of observed covariates. In the context of this study, it represents the likelihood of a student's participation in the Mathematical Thinking Workshops, based on their pre-university Mathematics and English performance. (Joffe & Rosenbaum, 1999; Austin, 2011a).

Propensity Score Matching (PSM): Propensity Score Matching is a statistical technique used in quasi-experimental research to reduce selection bias and estimate causal treatment effects in observational data. PSM creates matched pairs, or sets, of treatment and control students who have similar propensity scores, thereby approximating a randomised, experimental design. In this study, PSM is used to compare workshop participants with similar non-participants, thereby enabling a more accurate estimation of the Average Treatment Effect on the Treated (ATT).

National Senior Certificate (NSC): The National Senior Certificate (NSC) is the official school-leaving qualification in South Africa, awarded upon the successful completion of Grade 12. In this study, the NSC English and Mathematics scores of students were used as covariates to control for prior academic achievement when estimating the impact of the Mathematical Thinking Workshops.

National Benchmark Test (NBT): The National Benchmark Test (NBT) is a standardised assessment used in South Africa to evaluate students' academic readiness for university entrance. Developed by the Centre for Educational Testing for Access and Placement (CETAP) at the University of Cape Town, the NBT complements the National Senior Certificate (NSC) by assessing core academic competencies needed for success in higher education. In this study, students' NBT Mathematics and Quantitative Literacy scores were used as covariates.

Nearest-Neighbour Matching (with Replacement): Each treated unit is matched to the closest control unit based on a distance metric (here, logit-PS). "With replacement" allows

the same control to be used for multiple treated units, which improves match quality but induces dependence handled via clustering.

Treatment x Band Interaction: A regression term testing whether the treatment effect differs across levels of a moderator (here, NSC Mathematics bands: 70-84 vs 85-100). A significant interaction indicates heterogeneous effects by band.

Zone of Proximal Development (ZPD): The Zone of Proximal Development (ZPD) (Vygotsky, 1978) refers to the difference between what a learner can do independently and what they can do with the guidance of a more knowledgeable other (e.g., a teacher, peer, or facilitator). It represents the optimal space for learning, where tasks are just beyond the learner's current ability but achievable with appropriate support or scaffolding.[1]

[1] Although ZPD and scaffolding have become synonymous in the literature, they are, in fact distinct. Vygotsky never used the term scaffolding in his work; it was Wood et al. (1976) who introduced the term. Scaffolds may be used within the ZPD, but the ZPD happens over a long time, while scaffolds are task-specific.

LIST OF TABLES

Table 1: Comparative Perspectives on Identity in University Mathematics 24
Table 2: Comparative summary of ZPD, Cognitive Load Theory, and APOS (Adapted from Mokhithi et al.,(in press)). 31
Table 3: Overview of Mathematical Thinking Workshops, Mathematical Topics, and Metacognitive Skills 54
Table 4: Descriptive statistics of Cohort A and Cohort B. 94
Table 5: Absolute Standardised Mean Difference (ASMD) for covariates 98
Table 6: Mean covariates and assessment scores after caliper matching. 99
Table 7: ATT estimates by outcome with cluster-robust SEs 102
Table 8:Rosenbaum Sensitivity (Wilcoxon Signed-Rank) on matched pairs 103
Table 9: Covariate balance (ASMD) before and after overlap weighting. 105
Table 10: Workshop effects by pre-university mathematics achievement 105
Table 11: Average treatment effects on the treated (ATT) for higher-order (HO) items .. 107
Table 12: Higher-order (HO) performance by prior achievement 108
Table 13: Initial codes generated from the focus group interview transcripts. 116
Table 14: Nine preliminary themes and their codes. 120
Table 15: Refinement and grouping of preliminary themes into five final themes. 122
Table 16: Summary of the final themes and alignment with the research questions. 125
Table 17: Relationship between themes and research questions. 143

LIST OF FIGURES

Figure 1: APOS progression and key transitions (adapted from Villabona et al., 2024). .. 29
Figure 2: Workshop 3B, Question 6: Proof by Contradiction for the Irrationality (from Chamberlain and Vidakovic, 2021). 44
Figure 3: Proof by induction with varying levels of difficulty. 46
Figure 4: Question linking new concepts to existing mathematical understanding. 48
Figure 5: Workshop 1A: Jo Boaler insights on Lex Fridman's podcast 51
Figure 6: Worksheet 6 focused on pattern recognition in solving limits 52
Figure 7: An example of a preamble, Workshop 4A 53
Figure 8: Integrating case study and mixed methods designs (Guetterman & Fetters, 2018). 64
Figure 9: The Convergent Design (Creswell & Plano Clarke, 2017). 67
Figure 10:Proof from tutorials and past papers 78
Figure 11:Higher-order question for Squeeze Theorem 79
Figure 12: Propensity Scores overlap plot before Propensity Score Matching 95
Figure 13: Propensity scores overlap after NNM with caliper 97

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii
ABSTRACT iii
DEFINITION OF TERMS v
LIST OF TABLES ix
LIST OF FIGURES x
TABLE OF CONTENTS xi
CHAPTER 1: INTRODUCTION 1
1.1 Background and Context 1
1.2 Research Problem 3
1.3 Research Questions and Objectives 4
1.4 Significance of the Study 6
1.4.1 Theoretical contribution 6
1.4.2 Practical application 6
1.4.3 Methodological innovation 7
1.4.4 Contextual relevance and broader impact 7
1.5 Thesis Structure 7
1.5 Published Results 9
CHAPTER 2: LITERATURE REVIEW 10
2.1 Introduction and Purpose 10
2.2 Mathematical Thinking and Metacognition 11
2.3 Challenges in the Transition to University Mathematics 15
2.3.1 Inadequate foundational knowledge in mathematics 16
2.3.2 Cognitive Challenges: abstraction, definitions, and proof 18
2.3.3 Lack of metacognitive and self-regulated learning skills 20
2.4 Sociocultural and Identity Perspectives 22
2.5 Design Principles and Theoretical Foundations 25
2.5.1 ZPD and Intentional Mediation 25
2.5.2 Cognitive Load-Aware Sequencing 26
2.5.3 APOS-Oriented Concept Development 28
2.5.4 Integrative design rationale 30
2.6 Evaluating Educational Interventions 33

2.6.1 Introduction 33
2.6.2 Challenges in Evaluation 34
2.6.3 Methods for Evaluating Interventions 35
2.6.4 Conclusion for Evaluating Educational Interventions 36
2.7 Gaps in the Literature 37
2.7.1 Limited Focus on Metacognition and Mathematical Thinking 37
2.7.2 Lack of Theoretically Grounded Interventions 37
2.7.3 Limited Mixed-Methods Evaluations 38
2.8 Conclusion of the Literature Review 39
CHAPTER 3: WORKSHOP DESIGN AND IMPLEMENTATION 41
3.1 Introduction 41
3.2 Theoretical Foundations 41
3.2.1 Constructivism 41
3.2.2 APOS theory (design implications and concrete use) 43
3.2.3 The Zone of Proximal Development (ZPD) 45
3.2.4 Cognitive Load Theory 47
3.2.5 Theoretical Foundations Summary 49
3.3 Workshop Design 49
3.3.1 Workshop Structure and Format 49
3.3.2 Workshop worksheets and topics 51
3.4 Implementation Strategy 55
3.4.1 Recruitment of students 55
3.4.2 Role of Facilitators and Support Staff 55
3.4.3 Feedback and Improvement 55
3.5 Challenges 56
3.5.1 Timetables 56
3.5.2 Engagement, Motivation, and Adapting to Student Diversity 56
3.5.3 Coordination with the Mainstream Curriculum 57
3.5.4 Student Workload and Competing Priorities 57
3.5.5 Drop-in/Drop-out Participation 57
3.5.6 Stigma and Perception 57
CHAPTER 4: METHODOLOGY 59
4.1 Introduction 59

4.2 Research Paradigm (Pragmatic Worldview) 60
4.2.1 Role of the Researcher 61
4.3 Research Design 62
4.3.1 Quasi-Experimental Design 62
4.3.2 Case-Study Mixed-Methods Approach 63
4.3.3 Convergent Parallel Design 66
4.3.4 Case Study Context and Setting 67
4.4 Sampling and Participants 68
4.5 Data Collection Methods 71
4.5.1 Quantitative Data Collection 71
4.5.2 Qualitative Data Collection 72
4.6 Data Analysis 73
4.6.1 Quantitative Data Analysis 73
4.6.2 Qualitative Data Analysis 79
4.6.3 Integration of Quantitative and Qualitative Data 84
4.7 Validity and Trustworthiness 87
4.8 Limitations of the Research Design 88
4.9 Ethical Considerations 89
4.9.1 Informed Consent 89
4.9.2 Confidentiality and Anonymity 89
4.9.3 Voluntary Participation 90
4.9.4 Ethical Approval 90
4.10 Methodology Conclusion 90
CHAPTER 5: QUANTITATIVE RESULTS AND ANALYSIS 92
5.1 Introduction 92
5.2 Descriptive Statistics and Sample Construction 93
5.3 Propensity Score Estimation, Overlap, Matching, and Balance 95
5.3.1 Propensity Scores Estimation 95
5.3.2 Propensity Score Matching 96
5.4 Effect of Workshops on Academic Performance 100
5.5 Variation Across Student Sub-Groups 104
5.6 Performance on Higher-Order Questions 106
5.7 Higher-Order Questions (Variation Across Sub-Groups) 108

5.8 Summary of Quantitative Findings 110
CHAPTER 6: QUALITATIVE FINDINGS AND ANALYSIS 112
6.1 Introduction 112
6.2 Overview of Thematic Analysis Process 113
6.2.1 Familiarisation with the Data 114
6.2.2 Generating Initial Codes 115
6.2.3 Searching for and Reviewing Themes 119
6.2.4 Reviewing Potential Themes 121
6.2.5 Defining and Naming Themes 124
6.3 Emergent Themes (Producing the Report) 127
6.3.1 Theme 1: Developing Reflective and Strategic Mathematical Thinking 127
6.3.2 Theme 2: Growth in Mathematical Communication and Problem-Solving Confidence 130
6.3.3 Theme 3: Reconnecting with Mathematics Through Alignment with Coursework 133
6.3.4 Theme 4: Facilitator Influence and Supportive Learning Environment 135
6.3.5 Theme 5: Student Voices on Workshop Hurdles 139
6.4 Summary of Qualitative Findings 142
CHAPTER 7: DISCUSSION AND CONCLUSIONS 145
7.1 Purpose and Roadmap 145
7.2 Integrated Synthesis of Findings 145
7.2.1 Areas of Convergence 146
7.2.2 Areas of Divergence 147
7.2.3 Added Value of Mixed Methods Integration 148
7.3 Theoretical Integration and Mechanisms of Impact 149
7.3.1 Constructivist Foundations and Active Learning 150
7.3.2 APOS Theory and Cognitive Development 151
7.3.3 Zone of Proximal Development and Peer Collaboration 153
7.3.4 Belonging and the Affective Domain 155
7.3.5 Implications for Transitional Mathematics Programs 156
7.3.6 Contribution to Tertiary Mathematics Education 157
7.4 Practical Implications 160
7.4.1 Curriculum Alignment and Student Support Services 160

7.4.2 Scalability and Sustainability 163
7.5 Methodological Reflections and Study Limitations 165
7.5.1 Strengths of the Mixed Methods Design 165
7.5.2 Limitations and Constraints 167
7.5.3 Reflection on Data Collection and Analysis 170
7.5.4 Reflexivity and Researcher Positionality 172
7.6 Answers to the Research Questions 174
7.6.1 Quantitative Research Question 1 174
7.6.2 Quantitative Research Question 2 177
7.6.3 Quantitative Research Question 3 179
7.6.4 Qualitative Research Question 1 180
7.6.5 Qualitative Research Question 2 181
7.6.6 Qualitative Research Question 3 183
7.7 Contributions to Knowledge 185
7.7.1 Theoretical Contributions 185
7.7.2 Methodological Contributions 186
7.7.3 Practical Contributions 187
7.8 Recommendations 187
7.8.1 Recommendations for Workshop Design and Facilitation 188
7.8.2 For Educational Practice 190
7.8.3 For Institutional Policy 191
7.8.4 For Future Students 193
7.9 Directions for Future Research 193
7.9.1 Longitudinal Impact and Academic Trajectories 194
7.9.2 Workshop Design and Delivery Modes 194
7.9.3 Affective Dimensions and Identity Transformation 194
7.9.4 Institutional Implementation and Cross-Context Transferability 195
7.9.5 Artificial Intelligence and the Changing Educational Landscape 195
7.9.6 Methodological Innovation and Mixed Methods Evaluation 195
7.10 Final Reflections 196
REFERENCES 198
APPENDICES 228
Appendix A: Ethical Approval 228

Appendix B: Participant Information and Consent Forms 230
Appendix C: Focus Group Protocols (interview question sets) 231
Appendix D: Sample Focus Group Transcripts 234
Appendix E: Initial Codes 255
Appendix F: Sample Workshop Materials (Worksheet for Workshop 1A). 268
Appendix G: Higher Order Questions 275
Appendix H: Python Codes for Quantitative Analysis 282
H.1 Propensity Score Estimation 282
H.2 Descriptive statistics of Cohort A and Cohort B 283
H.3 Propensity Score Matching with Caliper and ATTs 284
H.4 Sensitivity Analysis 285
H.5 Variation Across Subgroups 286
H.6 Performance on Higher-Order Questions 289
H.7 Higher-Order Questions (Variation Across Sub-Groups) 290

CHAPTER 1: INTRODUCTION

1.1 Background and Context

The transition from school to university mathematics is widely recognised as one of the most challenging academic shifts students face, even for those who have excelled in school-level Mathematics (Lowe & Cook, 2003). At its core, this transition requires a cognitive reorientation, where students must shift from predominantly procedural, algorithmic approaches to more abstract, proof-based reasoning (Tall, 1997; Thomas et al., 2015). This cognitive leap is often accompanied by affective and identity-related struggles, including reduced confidence, heightened anxiety, and feelings of alienation from the mathematical community (Di Martino et al., 2023; Geisler & Rolka, 2021). Unsurprisingly, these combined pressures contribute to high rates of failure and attrition in STEM programmes globally. Such patterns are well documented in the UK, USA, and Australia, where institutions have responded with a range of targeted support interventions (Clark & Lovric, 2009; Pampaka et al., 2016; Penazzi et al., 2022).

In South Africa, the transition is further complicated by systemic inequalities and curriculum misalignments, which disproportionately affect students from historically under-resourced schools (Engelbrecht & Harding, 2015; Schöer et al., 2010). Although South African universities have developed various initiatives, such as foundational and augmented courses, supplementary tutorials, bridging courses, extended curriculum models, and peer mentoring programmes, research indicates that these primarily reinforce content knowledge (Boughey & McKenna, 2016; Council on Higher Education, 2020; Lewin & Mawoyo, 2014; Mouton & Rewitzky, 2025). While valuable, such approaches rarely address the deeper conceptual, metacognitive, and identity-related dimensions that are critical for sustained success in university mathematics (Ogude et al., 2012; Ogude & Rollnick, 2022). Moreover, many interventions lack explicit theoretical grounding, which makes their mechanisms of impact difficult to evaluate, refine, or replicate (Garraway & Bozalek, 2019).

This study evaluates the effectiveness and limitations of Mathematical Thinking Workshops (MTWs), a first-year university mathematics support programme deliberately designed to extend beyond content remediation. The workshops explicitly target students' mathematical reasoning, abstraction, and metacognition. Their design draws on Constructivism (von Glasersfeld, 1995), APOS theory (Asiala et al., 1996), Vygotsky's Zone of Proximal Development (ZPD) (Vygotsky, 1978), and Cognitive Load Theory (CLT) (Sweller, 1988). Each framework contributes a distinctive lens: Constructivism emphasises active knowledge construction; APOS theory describes how learners move from actions to processes, objects, and schemas; ZPD highlights the role of guided peer interaction and mediation; and CLT informs the careful sequencing of tasks to balance cognitive demand with learning potential.

The urgency of such an intervention is underscored by outcomes in the first-year mathematics course at the University of Cape Town, where this study is situated. Between 2015 and 2017, approximately one-third of students enrolled achieved below 50%, while a further 13.4% scored between 50% and 54% (Shay et al., 2020). More recent institutional data (Department of Mathematics and Applied Mathematics, UCT, 2025) show that in 2018, 33.8% of students scored below 50% and a further 23.0% between 50–59%; in 2019, 21.0% scored below 50% and 25.7% between 50–59%; and in 2022, 63.8% fell below 50%. The 2020 and 2021 cohorts are not included here, as the COVID-19 pandemic substantially altered pedagogical approaches, with most teaching and assessments taking place online. Notably, many students in this cohort completed portions of their pre-university schooling during the COVID-19 pandemic, with prolonged disruptions to teaching and assessment. This likely magnified gaps in preparedness relevant to the university transition. Although the headline pass rate often exceeded 65%, the quality of most of those passes is poor. Given that content-remediation supports are already in place at the University of Cape Town, these persistent outcomes provide a strong justification for a different intervention, one that foregrounds higher-order reasoning and metacognitive regulation. Moreover, while content-focused support can improve pass rates, the MTWs deliberately target conceptual reasoning and self-regulation that underwrite durable progress in university mathematics.

1.2 Research Problem

Although the transition to university mathematics has been widely studied (Di Martino et al., 2023), there is a lack of knowledge about interventions that specifically target the higher-order reasoning and self-regulatory skills essential for long-term success. Much of the existing research describes the difficulties students face, but fewer studies critically examine how structured support can strengthen mathematical thinking rather than bolster procedural fluency.

In South Africa, this lack of empirical insight is compounded by the diversity of students' prior experiences with mathematics. While interventions such as extended curricula and bridging programmes are well established, their design and evaluation have often prioritised content coverage. As a result, the field still lacks clear models of support that integrate conceptual development, metacognition, and identity formation in a way that is contextually relevant.

Methodologically, a further problem is that evaluations of support initiatives tend to privilege a single lens. Quantitative studies may indicate performance gains without clarifying the processes that drive them, while qualitative accounts may highlight valuable shifts in confidence or participation but remain disconnected from measurable outcomes. Without integrative approaches, it is difficult to draw robust conclusions about both the effectiveness of interventions and the mechanisms through which they work.

This investigation aims to fill existing research gaps by examining a theoretically grounded intervention, known as Mathematical Thinking Workshops (MTWs), through a convergent mixed-methods approach. The problem, therefore, extends beyond students' difficulties during the transition to university mathematics; it encompasses the field's deficiency in robust, theoretically informed evidence regarding how interventions can develop higher-order thinking and metacognitive regulation essential for ongoing success.

1.3 Research Questions and Objectives

This study aims to evaluate the effectiveness of MTWs in supporting first-year students' transition to university mathematics. The overarching research question guiding the study is:

How do Mathematical Thinking Workshops influence first-year Mathematics students' understanding and their ability to apply mathematical concepts?

This central question involves both cognitive and emotional aspects of learning, questioning not only whether the workshops enhance academic results but also how students perceive and interpret them. It focuses on changes in problem-solving techniques, metacognitive regulation, and students' broader engagement with university-level Mathematics.

To address the overarching question comprehensively, the study pursues the following sub-questions:

Quantitative Research Questions

1. To what extent do Mathematical Thinking Workshops impact students' academic performance when controlling for pre-university academic achievement?
2. How does the impact of the workshops vary across different levels of pre-university Mathematics achievement?
3. To what extent do the Mathematical Thinking Workshops impact students' performance on higher-order assessment items in first-year university mathematics?

Qualitative Research Questions

1. How do students perceive the impact of the workshops on their learning approaches and understanding of university mathematics?

2. What aspects of the workshops do students identify as most beneficial for their mathematical development?
3. What challenges and barriers do students experience in engaging with the workshop content and implementing the learnt strategies?

Objectives of the Study

1. To design and implement Mathematical Thinking Workshops grounded in Constructivism, APOS theory, the Zone of Proximal Development, and Cognitive Load Theory.
2. To evaluate the workshops' impact on academic performance using propensity score methods that account for pre-university achievement.
3. To explore students' experiences and perceptions of the workshops through focus group discussions.
4. To identify the workshop elements students found most beneficial and the challenges that constrained their learning.
5. To integrate quantitative and qualitative evidence to provide a comprehensive account of workshop effectiveness and the mechanisms through which impact occurs.

Objective 1 provides the foundation for all RQs by establishing the theoretically grounded intervention. Objective 2 aligns with Quantitative RQs 1–2 (estimating overall impact on academic performance and heterogeneity by pre-university achievement). Objective 3 aligns with Qualitative RQs 1–3 (students' learning approaches/understanding, perceived benefits, and experienced barriers). Objective 4 aligns with Qualitative RQs 2–3 (beneficial elements and constraints). Objective 5 aligns with Quantitative RQs 1–2 and Qualitative RQs 1–3 by integrating findings to explain effectiveness and mechanisms.

1.4 Significance of the Study

This study contributes to both the theoretical understanding and practical applications of supporting students' transition from high school to university Mathematics. Its significance lies in four key areas:

1.4.1 Theoretical contribution

The study demonstrates how Constructivism, the Zone of Proximal Development (ZPD), APOS theory, and Cognitive Load Theory (CLT) can be integrated into a single intervention in a South African context. It articulates design principles, such as task sequencing, scaffolding, and APOS-style encapsulation, that demonstrate how these frameworks can complement one another in practice. At the same time, it acknowledges that these lenses did not wholly drive the design: much of the initial pedagogy was guided by practitioners' intuition and iterative problem-solving, with the theories often applied as a reflective scaffold to rationalise, refine, and stress-test design choices. The study also interrogates limitations, for example, how APOS and ZPD operate when students have uneven prior preparation or limited experience with collaborative learning. Thereby contributing to debates about the applicability and interplay of learning theories in mathematics education and the productive tension between intuition-led design and theory-informed justification.

1.4.2 Practical application

The evaluation provides evidence-based insights for Mathematics departments and student support units seeking to cultivate first-year learning environments where students participate, persist, and feel a sense of belonging within the Mathematics field. It identifies facilitation practices that create low-threat learning spaces, collaborative structures that align with course demands, and design features that explicitly foster higher-order reasoning, metacognitive regulation, and community support. These findings can inform the development of sustainable, replicable models of support that extend beyond content remediation.

1.4.3 Methodological innovation

The study offers a comprehensive evaluation model that combines advanced quantitative methods with qualitative inquiry. By integrating propensity score matching, cluster-robust inference, Holm-adjusted p-values, and Rosenbaum sensitivity checks with reflexive thematic analysis, this approach illustrates how mixed methods can capture both the effectiveness of an intervention and the mechanisms through which its impact occurs. This methodological contribution responds to the field's need for evaluation designs that move beyond single-lens approaches.

1.4.4 Contextual relevance and broader impact

In South Africa, where educational inequality remains a challenge, it is important to understand effective intervention strategies for promoting equitable access to and success in higher education mathematics. The implications of this study may extend to national policy on foundational support, curriculum alignment, and equity in STEM education. Its findings identify a subgroup, consisting of students in the 70–84 NSC Mathematics band, for whom support may be especially beneficial, thereby offering evidence to guide targeted provision. Beyond the institutional level, the results may carry implications curriculum alignment, and equitable participation in STEM fields.

1.5 Thesis Structure

This thesis is organised into seven chapters, each of which addresses a distinct aspect of the research. Chapter 1 introduces the study's context, outlines the research problem, presents the research questions, and explains the significance of the study. Chapter 2 follows with a review of the relevant literature on the transition from school to university mathematics, the development of mathematical thinking and metacognition, identity and affect in mathematics learning, and the evaluation of educational interventions. This review highlights key gaps that the present study seeks to address.

Chapter 3 presents the design and implementation of the MTWs, outlining their theoretical grounding, workshop structure, worksheet design, facilitation practices, and the challenges encountered in implementation.

Chapter 4 then presents the mixed-methods research design, describing the research paradigm, data collection, and analytic strategies. It explains how propensity score methods were applied in the quantitative strand, how reflexive thematic analysis was employed in the qualitative strand, and how the two strands were integrated. Considerations of validity, trustworthiness, and ethics are also discussed.

The quantitative results are reported in Chapter 5, which presents findings on the academic impact of the workshops. This includes propensity score estimation, covariate balance, treatment effects on overall performance, subgroup variation by prior achievement, and outcomes on higher-order assessment items. Chapter 6 complements this analysis by examining qualitative data from focus group discussions, identifying themes related to students' development of mathematical thinking, metacognitive regulation, confidence, sense of belonging, and the challenges they encountered while engaging with the workshops.

Finally, Chapter 7 integrates the quantitative and qualitative findings to provide a comprehensive account of workshop impact. It interprets the results in relation to theoretical mechanisms, discusses practical implications for teaching and support, reflects on methodological strengths and limitations, and situates the study within the broader field of tertiary mathematics education. The chapter ends by addressing the research questions, highlighting the study's contributions to knowledge, and providing recommendations for future research and practice.

1.5 Published Results

Most of the qualitative findings reported in this thesis will be published in an international peer-reviewed journal article (the paper is already accepted), **“I Call It Math Therapy”: Student Narratives of Growth, Belonging, and Confidence in Mathematical Thinking Workshops** (Mokhithi et al., 2025). In this thesis, those findings are elaborated upon and integrated with quantitative analyses to provide a more comprehensive account than that offered in the published article.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction and Purpose

This chapter develops the literature-based rationale for the Mathematical Thinking Workshops (MTWs) and specifies how they are expected to work, for whom, and under what conditions. It begins by clarifying what is meant by mathematical thinking, a multifaceted capacity that encompasses problem-solving, conjecturing, explaining, abstracting, and proving, all of which are supported by metacognitive regulation. It then situates this capacity within the demands of first-year university mathematics, where formal definitions, higher levels of abstraction, and proof-based reasoning become central. The review takes into account the South African context of heterogeneous prior preparation, multilingual classrooms, and uneven access to high-quality mathematics teaching, treating these conditions as factors that shape opportunity to learn.

The chapter then assembles an integrated theoretical lens across four strands that recur throughout the thesis. From the cognitive strand, cognitive constructivism, APOS theory, and Cognitive Load Theory (CLT) inform how concepts are introduced, coordinated, and consolidated (e.g., from actions to processes, objects, and schemas; with extraneous load reduced and intrinsic complexity paced). From the metacognitive/Self-Regulated Learning (SRL) strand, planning, monitoring, and evaluation are positioned as teachable practices that enable transfer beyond familiar problem types. From the sociocultural/identity strand, ZPD-based mediation, participation structures, and recognition are treated as core design elements rather than add-ons, because identity is negotiated in interaction, and students must be positioned and see themselves as legitimate contributors. From the affect/motivation strand, the chapter links belonging, self-efficacy, and anxiety to participation and persistence.

Building on these strands, the chapter articulates design principles for the MTWs: calibrated challenge within the ZPD; APOS-oriented tasks that prompt coordination across representations, reversibility (working backwards to reconstruct the generating

action), and generalisation; load-aware materials and sequencing; facilitation and norms that aim to amplify student voice and distribute epistemic authority; and explicit metacognitive routines. A detailed account of how each theory informs specific worksheet tasks, activity sequences, and facilitation routines is provided in Chapter 3.

Finally, the chapter identifies gaps in the literature, notably the limited attention to metacognition in transition interventions, the frequent absence of explicit theoretical grounding, and the underutilisation of mixed-methods designs in the South African context, and positions the present study to address these gaps.

2.2 Mathematical Thinking and Metacognition

There is no universally accepted definition of mathematical thinking. It is widely recognised as a multifaceted construct that encompasses problem-solving, conjecturing, logical reasoning, proving, generalisation, and abstraction to explore and understand mathematical concepts and relationships (Breen & O'Shea, 2010; Mason et al., 1985). Leron (2003) argues that mathematical thinking extends beyond formal mathematics and serves as a way of thinking that is applicable across various disciplines and everyday situations. He further suggests that it enables individuals to analyse information, identify patterns, and construct logical arguments. In the context of this study, mathematical thinking refers specifically to the ability to systematically approach unfamiliar problems, reason through complex relationships, and construct meaningful solutions using mathematical representations and structures. It embraces both cognitive and metacognitive dimensions.

In contrast to school-level mathematics, university mathematics demands a more advanced form of thinking, where formal definitions, abstract reasoning, and proof-based logic are central (Hochmuth et al., 2021). Students are expected to move beyond a procedural understanding to develop conceptual depth and the ability to construct rigorous arguments (Tall, 1991). This shift can be overwhelming, as many students

struggle to adapt to the increased abstraction and the formal, axiomatic approach adopted in university courses (Engelbrecht, 2010; Selden & Selden, 1995).

Furthermore, in an international secondary school context outside South Africa, Vollstedt et al. (2014) found that “proofs are under-emphasised” (p. 47). Their work, together with other international studies, point to a broader pattern in which many students encounter relatively little to formal work with proof before entering university. In a study involving 535 early undergraduate students across six universities, Stylianou et al. (2015) observed that students often struggle to construct and comprehend mathematical proofs, despite recognising their importance in mathematics. Although these studies are situated outside South Africa, they resonate with South African research on the school-university transition, which reports that many entrants have had limited prior experience with proof-based reasoning and abstract argument (Engelbrecht & Harding, 2015; Msomi & Rzyankina, 2024), even though the extent of exposure varies across school types and regions. This challenge is compounded by the cognitive demands of university mathematics, which require students to integrate symbolic, graphical, and algebraic representations in a more interconnected and coherent manner (Swidan & Yerushalmy, 2016; Tall, 1991). The MTWs, therefore, include tasks, where possible, that move deliberately between these representations to surface underlying structure.

The cognitive demands imposed on students during their transition to university mathematics can also increase frustration and anxiety (Neumann et al., 2021). Working with abstract concepts and engaging in higher-order reasoning can create a sense of alienation, as many students feel unprepared for the level of thinking required (Mann, 2001; Pampaka et al., 2016). According to Geisler and Rolka (2021), this shift is not only a change in content but also a fundamental transformation in how students are expected to engage with mathematics. If this transition is not supported effectively, it could impede their progress. Consequently, addressing affect alongside cognition is fundamental to Mathematical Thinking Workshops.

Metacognition, often described as thinking about thinking, refers to an individual's awareness and ability to regulate their own cognitive processes in learning contexts (Alias & Sulaiman, 2017; Rhodes, 2019). It involves the ability to plan, monitor, and evaluate one's learning strategies and adjust them as necessary. Metacognition primarily encompasses two components, namely metacognitive knowledge and metacognitive regulation (Brown, 1987). Metacognitive knowledge refers to a learner's awareness of different types of tasks, the strategies available to approach them, and the learner's own cognitive strengths and limitations (Flavell, 1979). For example, learners with practical metacognitive knowledge skills are aware of the topics they have mastered and those they struggle with, and know different problem-solving strategies and when and how to use them. Metacognitive regulation, on the other hand, refers to strategies learners use in planning, managing, monitoring, and evaluating their cognitive processes (Schraw & Dennison, 1994), and for example, selecting a learning strategy for a task and realising that the strategy you are using to solve a mathematical problem is not working and trying another approach. Metacognitive competencies in learning, which are often absent from most course content (Sindhwani & Sharma, 2013), need to be explicitly taught; therefore, the workshops incorporate deliberate prompts for planning, monitoring, and reflection.

Mathematical thinking and metacognition are a central focus of the MTWs. These workshops aim to enhance students' ability to think mathematically by providing them with opportunities to solve open-ended problems collaboratively in groups and use multiple representations. Open-ended problem-solving tasks encourage students to think creatively and flexibly, while collaborative learning activities allow them to discuss and justify their reasoning with peers (Albert & Kim, 2013; Chan & Clarke, 2017). Multiple representations, including visual, symbolic, and verbal representations, alone do not deepen conceptual understanding; they require instructional support to be effective (Berthold & Renkl, 2009).

Research suggests that higher levels of metacognition are associated with better mathematical problem-solving performance (Izzati & Mahmudi, 2018). Metacognitive awareness enhances students' ability to analyse problems, plan solution strategies, and

evaluate their approaches. Metacognition is closely linked to self-regulated learning, which involves setting goals, managing time, seeking help, and persisting in the face of challenges (Zimmerman, 2002).

One of the workshop's goals is to enhance students' metacognitive skills by promoting reflection, encouraging self-regulation, and fostering a growth mindset[2]. Reflection activities, such as group discussions, offer students opportunities to reflect on their problem-solving processes and identify areas for improvement (Mason & Singh, 2010, 2016). Self-regulation skills, such as setting goals and monitoring progress, help students manage their learning effectively (Cera et al., 2013). At the same time, a growth mindset motivates students to perceive challenges as opportunities for learning and growth, rather than as threats to their competence (Dweck, 2006).

Mathematical thinking and metacognition are deeply interconnected; a student's capacity for mathematical thinking is closely linked to their ability to regulate cognitive processes (Stanton et al., 2021). By applying metacognitive strategies, students can evaluate the effectiveness of different problem-solving approaches and make necessary adjustments, thereby enhancing their mathematical reasoning (Schoenfeld, 2009). The bidirectional relationship between mathematical thinking and metacognition highlights how metacognitive skills enhance mathematical problem-solving (Izzati & Mahmudi, 2018). Simultaneously, engaging in complex mathematical tasks fosters the development of those very metacognitive abilities (Kim & Moore, 2019). When solving challenging problems, students must reflect on their thinking, test hypotheses, and adapt their strategies practices that further strengthen metacognitive abilities (Kapa, 2001). Learners who develop strong metacognitive skills, such as task analysis, planning, monitoring, and reflection, tend to achieve better results in mathematics (Noor, 2022; Tachie, 2019).

The MTWs are explicitly designed to foster both mathematical thinking and metacognitive skills. Problem-solving tasks require students to apply mathematical thinking skills as they

[2] A growth mindset is the belief that a person's capacities and talents can be improved over time (Dweck, 2006).

plan, monitor, and reflect on their strategies (Stillman, 2011). Group activities encourage learners to articulate their reasoning, evaluate alternative approaches, and deepen their understanding of concepts. Reflective exercises, such as group discussions, help students develop metacognitive awareness and self-regulation. Through these structured activities, the workshops intentionally aim to cultivate students' mathematical thinking and metacognitive skills.

Significantly, the workshops' approach extends beyond mathematical thinking and metacognition to include multiple psycho-social and study skills, belonging and identity, self-efficacy, help-seeking, and time management.

2.3 Challenges in the Transition to University Mathematics

Students worldwide face a range of academic and affective challenges during the transition to university-level mathematics. In South Africa, these challenges are exacerbated by deep-rooted educational inequalities and resource disparities, which disproportionately affect students based on geographic location, school type, and socio-economic background. Students from rural or underfunded township[3] schools often enter university with more pronounced foundational gaps than their peers from urban, well-resourced schools. These disparities underscore the need for designing support interventions that are tailored to the diverse academic backgrounds of students.

This section synthesises the following persistent challenges associated with the transition to university mathematics:

a. Inadequate foundational knowledge in mathematics
b. Cognitive challenges
c. Metacognitive/SRL challenges

[3] In South Africa, a township is a historically segregated residential area, often underdeveloped, that was primarily designated for non-white populations (Black Africans, Coloureds, and Indians) during the apartheid era. These areas were typically located on the outskirts of towns and cities, separate from the "white areas".

2.3.1 Inadequate foundational knowledge in mathematics

A persistent challenge in South Africa's education system is the inadequate mathematical foundation acquired by many learners (Jojo, 2019), which undermines their preparedness for the demands of university-level mathematics. This problem is deeply rooted in systemic issues within the broader educational framework. South African primary and secondary school mathematics and science education is internationally recognised as among the weakest (Mabena & Mokgosi, 2021; Ntsanwisi, 2024; Wolhuter, 2014), limiting learners' access to and success in higher education.

Gaps in foundational knowledge can be attributed to multiple, intersecting factors, including disparities in resource allocation (Lee, 2011), variations in teacher quality (Jojo, 2019), language barriers in the learning of mathematics (Botes & Mji, 2010), and entrenched socio-economic inequalities (Mensah & Baidoo-Anu, 2022). These factors collectively contribute to a significant proportion of students entering university underprepared for the conceptual depth, abstract reasoning, and problem-solving required in mathematics-intensive programmes. The consequences are evident in high failure rates, low retention rates, and overall poor academic performance in STEM degree programmes.

A key contributor to inadequate foundational knowledge in mathematics is the unequal distribution of resources between historically disadvantaged schools and those in more affluent communities (Sedibe, 2011; Mokgwathi, 2025). Schools[4] in rural and township contexts often lack essential resources, including textbooks, adequately qualified mathematics teachers, and sufficient infrastructure (Du Plessis & Mestry, 2019; Masango et al., 2020). Such deficits severely limit the quality of mathematics instruction and learning opportunities available to learners. As a result, many students from these schools enter university with significant gaps in their mathematical knowledge and skills, making

[4] The enduring effects of apartheid-era policies, most notably the *Bantu Education* system introduced in 1953, continue to shape educational inequalities in South Africa. This policy deliberately provided inferior education to Black South Africans, restricting access to quality mathematics and science education and entrenching systemic disparities that disproportionately affect historically disadvantaged communities to this day (Christie & Collins, 1982; Fiske & Ladd, 2004).

it challenging to engage with the abstract, formal, and conceptually demanding nature of university mathematics. Spaull and Jansen (2019) research corroborates the severity of this crisis, demonstrating that the quality of education in South Africa remained profoundly unequal decades after the end of apartheid, with students from disadvantaged backgrounds disproportionately affected by these disparities.

These resource-related inequalities often intersect with linguistic challenges. In a multilingual country like South Africa, many students receive mathematics instruction in a language other than their native language, which can hinder their comprehension of complex concepts and the development of effective problem-solving strategies. Durand-Guerrier et al. (2015) highlight the difficulties students faced in multilingual classrooms, where the language of instruction can act as a significant barrier to meaningful engagement with mathematical content. This challenge is especially pronounced among students from rural or disadvantaged backgrounds who may have had limited exposure to English (Khoza-Shangase & Kalenga, 2024), yet must navigate university-level mathematics in an English-medium context.

Underlying both the resource and language dimensions is the broader influence of socio-economic inequality, which appears to significantly shape students' readiness for university mathematics. This readiness is often influenced by social and economic circumstances, many of which may be traced to structural inequalities rooted in apartheid-era policies. Learners from low-income households often attend under-resourced schools and face additional challenges, including limited access to tutoring, technological tools, and other academic support systems (Njeri & Taym, 2024; Setlhare et al., 2016). Böhmer and Wills (2025) demonstrate a strong correlation between socioeconomic status and academic achievement in South Africa, with students from more affluent backgrounds significantly more likely to succeed academically. This inequality, a direct legacy of apartheid's spatial and educational segregation, perpetuates cycles of disadvantage: students from low-income backgrounds frequently enter university with substantial gaps in foundational knowledge, placing them at a marked disadvantage compared to their peers.

The cumulative effect of these factors is reflected in persistently high failure rates in first-year university mathematics, as documented in recent South African reviews of the transition to university mathematics (Msomi & Rzyankina, 2024). Many students enter university struggling with core concepts such as algebra, functions, and problem-solving strategies and skills that form the basis for engaging with more advanced mathematical content (Nortvedt & Siqveland, 2018). In response, various interventions have been implemented to address these gaps, including bridging programmes and curricular reforms. One such initiative is the Extended Curriculum Programme (ECP), introduced in many South African universities to support underprepared students by strengthening foundational knowledge and fostering skills required for success in university mathematics (Ramaila et al., 2019). While ECPs have shown some positive outcomes, they often lack a robust theoretical grounding (Garraway & Bozalek, 2019). They are not always designed to address the specific needs and contexts of students from disadvantaged educational backgrounds (Ogude & Rollnick, 2022).

2.3.2 Cognitive Challenges: abstraction, definitions, and proof

The transition from school to university mathematics represents a significant cognitive and conceptual leap for students. It involves moving from concrete, computation-based tasks to abstract, theoretical engagement with mathematics (Tall, 1997; Engelbrecht, 2010). For many, this is among the most challenging aspects of the transition, as it demands a deeper conceptual understanding and fundamentally different ways of reasoning about mathematical ideas (Cullinane, 2011; Fatqurhohman, 2016). In South Africa, this challenge is compounded by weak foundational competencies and limited practice with non-routine, higher-order tasks that characterise many students' school experiences (Department of Basic Education, 2024; Reddy et al., 2022). Understanding the nature of these shifts is therefore critical for designing interventions, such as the MTWs (a tertiary education level intervention) examined in this thesis, that aim to actively support students in navigating this phase of their academic development.

At the school level, mathematics education is often characterised by a focus on procedural teaching and rote learning. Students are typically trained to follow step-by-step procedures to solve well-defined problems, with an emphasis on memorising formulas and algorithms. While this approach can build basic computational fluency, it does little to cultivate the conceptual understanding, abstraction, and problem-solving flexibility required for success in university mathematics. Tall (2008) argues that a procedural focus in mathematics education leaves students ill-prepared for the critical shift from asking *how* to asking *why*, a hallmark of higher mathematics, where the emphasis moves beyond mechanical equation-solving to engaging with underlying principles, constructing proofs, and generalising concepts to unfamiliar contexts.

A related and equally significant change is the move from concrete to abstract thinking (Tall, 1991). School mathematics often situates concepts within familiar, real-world applications, making them accessible, but this approach can also limit exposure to formal abstraction. On the other hand, university mathematics introduces students to highly abstract concepts such as limits, continuity, and formal proofs, which demand cognitive flexibility and a conceptual understanding that many students find challenging (Tall, 1997). Selden and Selden (2003) argue that this shift is particularly challenging for students who are not exposed to rigorous mathematical reasoning in school, as they often struggle to comprehend the formal definitions and logical structures that underpin university mathematics.

The nature of problem-solving also changes substantially. In school settings, problems are typically closed-ended and solvable by applying known algorithms to familiar problem types (Jäder et al., 2019). At university, however, students are often faced with open-ended problems where the solution path is not immediately apparent. Schoenfeld (2009) stresses the importance of metacognitive abilities in this process, noting that successful problem-solving requires students to monitor their thinking, reflect on their strategies, and adapt their approaches as necessary. However, it seems that many students enter university relying more on memorisation and repetition than on reflective and adaptive

thinking, which may leave them less prepared for the demands of a more complex problem-solving landscape.

In South Africa, students from disadvantaged backgrounds, in particular, often enter university with significant gaps in their foundational knowledge and limited exposure to higher-order thinking (Centre for Educational Assessments, 2024). This makes the transition to university mathematics even more daunting. Many schools lack qualified mathematics teachers; therefore, students are often taught by underqualified or inadequately trained teachers (Luneta, 2024), which further exacerbates the problem.

The consequences of these often-unaddressed cognitive and conceptual gaps may lead to high failure and low retention rates in first-year mathematics courses. As Scott et al. (2007) observe, the transition to university mathematics entails a substantial cognitive leap, necessitating students to develop new approaches to thinking and engaging with the subject. Without structured support, many students struggle to make a successful transition. The MTWs respond to this need by providing a learning environment that encourages students to engage in open-ended problem-solving, collaborative learning, and reflective practice, to develop the conceptual and metacognitive skills essential for success in university mathematics.

2.3.3 Lack of metacognitive and self-regulated learning skills

The transition from school to university mathematics requires students to confront new content and modes of reasoning while simultaneously adopting fundamentally different ways of engaging with the subject. In school, learning is often teacher-centred (Hardman, 2023), where students passively receive information and practice predetermined procedures. University mathematics, in contrast, requires students to take ownership of their own learning. This includes managing their time, seeking help when needed, and actively engaging with the material (Ramāna et al., 2019; Tularam, 2013). Students who are accustomed to a more structured and guided approach to learning may find this shift particularly challenging. Many students encounter difficulties adapting to the demands of

university mathematics and often struggle with the increased autonomy and self-directed learning required at the tertiary level (De Guzmán et al., 1998).

In mathematics, metacognitive skills enable students to approach problems systematically, monitor their problem-solving processes, and identify gaps in understanding (William & Maat, 2020). Self-regulated learning, in contrast, involves goal setting, time management, help-seeking behaviour, and persistence in the face of challenges (Vosniadou, 2020). These abilities are critical for success in university mathematics, where the workload is generally heavier than it was in high school, the content is often abstract, and the problems are open-ended. However, many students enter university without these skills, mainly because their school experiences emphasised rote learning and procedural knowledge over reflection, self-monitoring, and independent problem-solving (Geduld, 2019; Jojo, 2019).

This skills gap has clear consequences. Students lacking effective planning and time management often find themselves overwhelmed by the pace and demands of university mathematics courses (Taylor & Mander, 2003). Research consistently shows a strong link between academic success and self-regulated learning, with students who set clear goals, organise their study schedules, and monitor their progress more likely to succeed (Guntur & Purnomo, 2024; Lindner & Harris, 1992; Schunk, 1984; Theobald, 2021; Xie et al., 2024; Xu et al., 2022; Zimmerman, 1990). Likewise, the absence of metacognitive awareness can result in a superficial grasp of mathematical concepts, whereby students can perform calculations but struggle to apply their knowledge in novel or unfamiliar contexts. Pintrich (2000) emphasises that students who can reflect on their understanding and adjust their strategies are far more likely to achieve conceptual mastery.

Another barrier is students' reluctance to seek help, whether due to fear of appearing incompetent, low academic self-efficacy, impostor feelings (Chen & Son, 2024; Ryan et al., 1998), or simply a lack of awareness of available support. In some school contexts, overt questioning is informally discouraged – often perceived as challenging authority or disrespect, and teachers may also restrict open questioning to maintain authority or avoid

exposing gaps in their own understanding when they feel underprepared, further suppressing help-seeking (Buseri, 1987; Findlay, 1995). Karabenick and Dembo (2011) emphasise that effective help-seeking is a crucial component of self-regulated learning, enabling students to overcome difficulties and persist in their academic pursuits. Without these behaviours, students may endure prolonged struggles with complex concepts, leading to frustration, disengagement, and, in some cases, attrition (Johnson et al., 2022; Won & Chang, 2024).

Addressing the absence of metacognitive and self-regulated learning skills must therefore be an explicit focus of interventions designed to support the transition to university mathematics. The MTWs aim to contribute to this goal by engaging students in reflective practice, collaborative problem-solving, and goal-setting activities. These workshops not only help participants develop metacognitive awareness and self-regulated learning strategies but also foster help-seeking behaviours and effective use of academic resources, equipping them with the skills needed to thrive in the university mathematics environment.

2.4 Sociocultural and Identity Perspectives

According to Wenger (1998), learning mathematics at university entails joining a new community of practice (CoP) with unfamiliar norms for explanation, precision, and critique, where learning becomes a trajectory from legitimate peripheral participation (LPP) to full participation in a CoP. As Young and Bunting (2024, p. 8) describe LPP, ‘legitimate’ means offering newcomers authentic roles that build competence in the knowledge, activities, and discourses of the community. 'Peripheral' implies that learners occupy a less central yet visible and necessary position in the community, while 'participation' signals that learning occurs as newcomers engage in the meaningful practices of their new community (Bunting & Williams, 2017). In large, time-pressured courses, many students remain peripheral observers rather than contributors. The MTWs are structured to create regular, small-group spaces where legitimate participation is possible from the

outset through authentic roles, public reasoning, and shared repertoires of presentations and strategies.

Identity can be understood as “being recognised as a certain kind of person in a given context” (Gee, 2000, p.99). According to Gee (2000), recognition operates through Discourse (how people talk to or with the learner) and Affinity (membership in practice-based groups). Facilitation that invites conjecture, revoices tentative ideas, and treats errors as data broadens what counts as mathematical competence, thereby redistributing authority beyond the lecturer (Lo & Ruef, 2020). Conversely, norms that privilege speed, correctness, or solitary work can marginalise less confident students, narrowing opportunities to be recognised as capable contributors (Gee, 2000).

In Wenger’s (1998) terms, movement from LPP toward fuller membership unfolds through mutual engagement, joint enterprise, and a shared repertoire. MTWs enact these conditions by seating students in stable groups, assigning joint problem-solving roles (e.g., summariser, sceptic, scribe), and building standard tools and language across weeks. As students contribute increasingly sophisticated strategies, they accrue participatory capital that supports a stronger mathematics identity.

Sfard and Prusak (2005) conceptualise identity as a set of narratives, stories about who one is now (Actual identity) and who one can become (Designated identity). These stories are negotiated socially and then internalised. MTWs reflection prompts (e.g., Describe a moment you felt like a mathematician today) are used to surface and reshape deficit narratives (e.g., I always mess up proofs) into growth-oriented ones (e.g., I can develop proof strategies with practice), aligning Actual and Designated identities over time.

Table 1 synthesises these perspectives and shows how they inform both design and analysis in this study. The table compares Gee’s (2000) recognition lens, Wenger’s (1998) participation lens, and Sfard and Prusak’s (2005) narrative lens across their core view of identity, primary locus, mechanisms of formation, design implications, and evaluation cues.

Table 1: Comparative Perspectives on Identity in University Mathematics

	Gee (2000)	**Wenger (1998)**	**Sfard & Prusak (2005)**
Core view of identity	Being recognised as a certain kind of person in a given context	Identity as a trajectory of participation in a Community of Practice	Identity as a set of narratives (stories) about who one is and might become
Primary locus[5]	Recognition through Discourse and Affinity groups	Participation in shared practice; mutual engagement, joint enterprise, shared repertoire	Discursive accounts: Actual vs Designated identities
Mechanism of formation	Recognition is negotiated in discourse, by others and by the self, so that positions are either taken up or resisted	Movement from legitimate peripheral participation to fuller membership via practice	Identities are authored, told, and retold; narratives are negotiated socially, then internalised
What counts as evidence in workshops	Who gets uptake; who is described as good at maths; who speaks and is responded to	Who takes roles, contributes strategies, uses shared tools; evolution of participation.	How students tell their maths stories (e.g., from "I'm bad at proofs" to "I can craft proofs")
Design implications	Revoice tentative ideas; widen what counts; track who gets uptake	Small-group roles; routines that build a shared repertoire; pathways from LPP to fuller membership	Reflection prompts that reframe deficit narratives as growth narratives; a curated portfolio of identity stories
Evaluation cues in this study	Evidence of broadened recognition and equitable uptake	Increased role-taking, strategy-sharing, and	Shifts from Actual to Designated identity claims in

[5] Primary locus: the principal site where a lens locates identity, where it is constituted and most observable , and thus the most sensible focus for design and evaluation (e.g., recognition in interactions, participation in communities, or narratives about the learner).

		repertoire use over time	interviews/reflections

2.5 Design Principles and Theoretical Foundations

The MTWs integrate the cognitive, social, and affective dimensions of learning. From the cognitive side, cognitive constructivism[6], APOS theory, and Cognitive Load Theory (CLT) guide how concepts are introduced, coordinated, and consolidated (Arnon et al., 2014; Piaget, 1950; Sweller, 1988). From the sociocultural side, including social constructivism traditions, ZPD-based mediation and identity-aware participation structures ensure students have legitimate opportunities to contribute and to be recognised as mathematical thinkers (Daniels, 2001; Vygotsky, 1978). A detailed account of how each theory informs specific worksheet tasks and workshop activities/facilitation is provided in Chapter 3.

2.5.1 ZPD and Intentional Mediation

The Zone of Proximal Development (ZPD), introduced by Vygotsky (1978), is a developmental principle that highlights the role of social interaction and mediation in learning. It refers to the distance between what a learner can do independently and what they can achieve through mediated guidance from a more competent other, such as a teacher or peer (Morgan & Skaggs, 2016). Learning within the ZPD extends beyond task completion and involves mediated development, where cognitive and psychological growth are supported through sustained interaction and guidance (Daniels, 2001).

In the MTWs, ZPD is made practical by pitching tasks just beyond students' independent competence and then making that challenge accessible through facilitative questioning, public uptake of student ideas, and structured collaboration. The facilitator and tutors

[6] Piaget's work is best referred to as cognitive constructivism in that it focuses on the development of the learner's cognition. While Vygotsky's work very specifically harnesses the social and cultural realm's impact on development and asserts that learning precedes development, Piaget's suggests that cognitive development leads learning. Although these views may seem contradictory, both theorists provide foundational contributions to constructivism, each from a different perspective.

press for representation, connection, and justification asking, for instance, what a particular step achieves or how a diagram constrains an algebraic move so that reasoning advances without answers being supplied. Students' partial strategies are revoiced and probed until they become shared resources for the group, which distributes epistemic authority and reinforces recognition. Collaboration is organised so that learners alternate between offering and receiving support, allowing each participant to experience both sides of mediated learning.

Although scaffolding is often associated with the ZPD, the MTWs distinguish between scaffolding and mediation. Scaffolding refers to temporary, task-specific support, such as worked-example prompts or hint sequences. Mediation, on the other hand, denotes the sustained developmental support carried across sessions through norms, talk moves, and reflective routines (Daniels, 2001; Wood et al., 1976). In practice, scaffolds are gradually faded as students' repertoires grow, while mediation remains a constant feature of the learning environment.

A central implementation challenge is identifying each learner's zone and maintaining an optimal level of difficulty that is demanding enough to promote growth yet not so difficult that it induces disengagement. In the South African context, where heterogeneous schooling histories and language backgrounds are common, ZPD-based mediation also serves an affective purpose. By normalising uncertainty and making progress visible, the workshops aim to help counter isolation and impostor feelings that often discourage participation in university mathematics.

2.5.2 Cognitive Load-Aware Sequencing

CLT informs the micro-design of tasks and materials by shaping how information is presented, sequenced, and consolidated, on the premise that working memory is limited and easily overburdened (Chandler & Sweller, 1991; Cowan, 2001; Sweller, 1988).

Cognitive load is categorised into three types: intrinsic, extraneous, and germane (Orru & Longo, 2019). Intrinsic load arises from the inherent complexity of the content, specifically the number of interacting elements that must be coordinated for understanding (Sweller, 1988). Concepts vary in complexity, and the tasks that are higher up Bloom's Taxonomy of Learning (Anderson & Krathwohl, 2001) will require more intrinsic cognitive allocation. Consequently, certain concepts are inherently more complex and require a greater cognitive load to be understood (Clark & Kimmons, 2023). It is best managed by decomposing complex problems into coherent sub-goals and by pacing element interactivity through a progression from worked examples to completion tasks and then fully independent items (Sweller, 2011).

Extraneous load is generated by the manner of presentation rather than the content itself; it is reduced by clean layouts, economical notation, and integrated text-diagram formats that avoid split attention so that scarce resources are not squandered on avoidable processing (Chandler & Sweller, 1991). Germane load/resources refer to the portion of cognitive effort productively invested in building and refining schemas. Design can foster this by prompting compare–and–contrast analyses, asking students to justify why a chosen method is efficient, and highlighting structural invariants across problems (Sweller, 2011). Sweller (2011) distinguishes germane cognitive load from the rest:

> Unlike intrinsic and extraneous cognitive load, germane cognitive load does not constitute an independent source of cognitive load. It merely refers to the working memory resources available to deal with the element interactivity associated with intrinsic cognitive load. (p. 126)

While CLT offers powerful design guidance, the literature also notes important caveats. First, the status of germane load has been debated: some scholars argue it is not a separable category from intrinsic load (Schnotz & Kürschner, 2007), and later CLT formulations partially fold germane into intrinsic load, creating definitional instability that complicates design prescriptions (Sweller, 2011). Second, measurement remains challenging; many studies rely on subjective mental-effort ratings, raising validity

concerns and prompting calls for triangulation with performance and process data (Paas et al., 2003). Third, distinguishing between extraneous and intrinsic load can be challenging in authentic classroom/workshop tasks, where what appears extraneous may, in fact, be conceptually integral; critics warn against oversimplifying complex problems in ways that hinder conceptual growth (de Jong, 2010). Finally, the "expertise-reversal effect"[7], cautions that support for novices may hinder more advanced learners, requiring careful fading and adaptation over time (Kalyuga, 2007).

In the MTWs, supports are deliberately faded over time, early sessions make strategic thinking explicit with cues and prompts. In contrast, later sessions withdraw those cues to manage intrinsic and extraneous demands while steering germane effort toward durable schema construction and transfer (Sweller, 2011).

2.5.3 APOS-Oriented Concept Development

APOS theory is a constructivist account of how mathematical concepts are built as learners move from actions to processes, then to objects, and ultimately to integrated schemas (Arnon et al., 2014; Dubinsky, 1991; Dubinsky & McDonald, 2001). Rooted in Piaget's notion of reflective abstraction[8], APOS treats new knowledge as the coordination and reorganisation of prior mental actions into more powerful structures (Dubinsky, 1991; Piaget, 1950). Two transitions are pivotal in this trajectory. Through interiorisation, externally performed, step-by-step actions are transformed into a coherent mental process that can be reflected on and manipulated. Through encapsulation, this process becomes an object, a unit of thought that can itself be transformed and related to other objects. Afterwards, objects and relations consolidate into a schema that organises understanding across contexts (Figure 1).

[7] "The expertise reversal effect demonstrates that instructional techniques must be tailored to the learner's existing knowledge base to avoid redundant or counterproductive cognitive demands (Kalyuga, 2007, 2009)."(Vasilaki & Mavrogianni, 2025, p. 4)

[8] Reflective abstraction (after Piaget) is the construction of new mathematical structures by reflecting on and reorganizing one's own actions and operations—interiorizing them into processes, encapsulating them as objects, and integrating them into schemas—rather than merely extracting features from external objects (Piaget, 2001; Arnon et al., 2014).

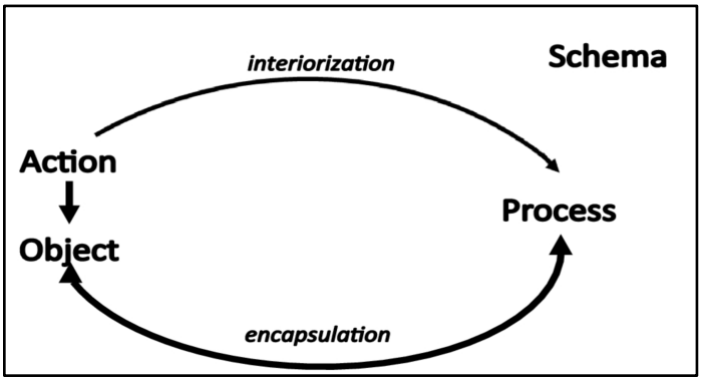


Figure 1: APOS progression and key transitions (adapted from Villabona et al., 2024).

In the MTWs task design, the progression is deliberately staged. Students first enact procedures (action), then articulate and coordinate those procedures mentally (process), and finally treat the process as an object that can be transformed, related, or generalised (object). To make these shifts visible, problems routinely require reversals of reasoning for example, reconstructing a generating action from a composed process, so that learners gain control over both forward and inverse moves. Facilitators press for encapsulation by asking students to name the process explicitly and then operate on it as a single entity (e.g., “If this limit rule is an object, what transformations are legitimate on it?”). Over time, these experiences accumulate into schemas as learners connect symbolic, graphical, numerical, and verbal representations and recognise invariant structure across topics (Arnon et al., 2014).

APOS does not imply a rigid, linear ladder. Learners often move back and forth between stages as their understanding deepens, and different sub-concepts within a topic may sit at different APOS levels simultaneously. To guide design, the study employs genetic decomposition, a hypothetical, epistemic model of the actions, processes, objects, and coordinating structures believed to underlie a target concept which is then used to craft and sequence tasks that are likely to elicit those structures (Trigueros & Oktaç, 2019). Because APOS is primarily cognitive, the workshops pair APOS-aligned tasks with ZPD-based mediation and a low-threat climate so that social participation and recognition support the very conceptual shifts the tasks are engineered to provoke. A detailed

mapping of how APOS informs specific worksheet items and activity sequences is presented in Chapter 3.

While APOS has been productive for analysing and designing task sequences, several cautions are noted in the literature. The genetic decomposition that anchors APOS is explicitly hypothetical and often assembled from researchers' experiences, textbook analyses, and classroom observations; this makes it susceptible to subjectivity and variability across studies, even when validation cycles are employed (Oktaç et al., 2019). Closely related is a common misreading of APOS as a linear stage model. Moreover, designing, testing, and refining robust genetic decompositions and turning them into coherent task sequences demands substantial pedagogical content knowledge and time, a challenge acknowledged in APOS-aligned task-design work (Trigueros & Oktaç, 2019).

2.5.4 Integrative design rationale

Building on the lenses developed in the preceding sections, the MTWs synthesise complementary theories to engineer a learning environment that is both conceptually demanding and socially supportive. APOS theory organises the intended conceptual trajectory from enacted procedures toward coordinated processes, encapsulated objects, and integrated schemas (Arnon et al., 2014; Dubinsky, 1991). ZPD-based mediation ensures that this progression is pursued at the edge of students' independent competence through sustained, dialogic support rather than answer-giving (Daniels, 2001; Vygotsky, 1978). Cognitive Load Theory informs the presentation and sequencing of tasks so that extraneous demands are pared back, intrinsic complexity is paced, and germane effort is channelled toward schema construction (Chandler & Sweller, 1991; Sweller, 1988). Table 2 maps each framework to specific pedagogical choices in the workshops and notes practical tensions in their application.

Table 2: Comparative summary of ZPD, Cognitive Load Theory, and APOS (Adapted from Mokhithi et al.,(in press)).

Category	**ZPD**	**CLT**	**APOS theory**
Primary Focus	Socially mediated learning through support	Mental efficiency during learning	Cognitive development in mathematical thinking
Schema Acquisition	Occurs via social interaction and scaffolding/mediation within the ZPD	Optimised by managing intrinsic, extraneous, and germane load	Emerges through progression from actions to processes to schemas
Role of the Learner	Active participant in a social context	A cognitive processor whose capacity must be managed	The learner progresses through internal mental constructions
Role of the Facilitator/Tutor	Provides scaffolding/mediation to bridge gaps in understanding	Design tasks that avoid overload and promote schema construction	Design activities that trigger transitions through APOS phases
Nature of Support	External and temporary; tailored to learner(s) readiness	Instructional design-driven; reduces unnecessary demands	Internalised through mathematical activity and reflection

Relevance to Workshops	Supports collaborative and peer-facilitated discussions	Informs task design to manage cognitive demand	Frames cognitive growth in mathematical understanding
Tensions/ Challenges in applying the theory	Identifying the learner(s)'s current zone and ensuring the support provided is neither overwhelming nor under-stimulating the learner(s).	Striking the right balance between reducing extraneous load and maintaining sufficient task complexity to promote meaningful learning.	Moving from theoretical genetic decomposition to practical task design requires a deep understanding of pedagogical content knowledge and iterative refinement.

The details of how APOS theory, ZPD, and CLT shaped the design of the workshops will be set out in Chapter 3

The combination of these theories also reflects a deliberate, pragmatic eclecticism. Although ZPD, APOS, and CLT originate from different intellectual traditions, with distinct ontological and epistemological commitments, their selective integration is justified by the multifaceted nature of learning in first-year mathematics. Each framework illuminates a different dimension of the learning process: ZPD foregrounds social mediation, APOS provides an account of the development of mental structures, and CLT focuses on managing cognitive resources during problem solving. In a practice-oriented intervention such as the MTWs, no single theory sufficiently captures the interplay of conceptual, cognitive, and social processes involved in mathematical thinking. However, combining these perspectives also entails recognising their limitations. CLT's information-processing assumptions are not fully commensurable with the sociocultural orientation of ZPD, and APOS theory's genetic decompositions do not always align neatly with the dynamic scaffolding implied by Vygotskian mediation. The MTWs therefore adopt these lenses instrumentally for both the design of workshop activities and the interpretation of student learning processes, drawing on their complementary strengths while remaining transparent about the conceptual tensions and constraints involved.

2.6 Evaluating Educational Interventions

2.6.1 Introduction

When evaluating educational interventions, it is essential to determine what works, for whom, and under what conditions. Effective evaluation should demonstrate not only whether an intervention improves outcomes, but also how and why it works, which is crucial for both research and policy. Researchers are encouraged to systematically evaluate interventions using appropriate research methods that support causal inference, ensuring interventions are evidence-based, scalable, and have a positive social impact (Slavin, 2017; Schlotter et al., 2011).

The primary aim of this study is to assess the impact of MTWs on students' academic performance and their cognitive and metacognitive development. This section provides a

review of the methodological approaches that can be used to evaluate educational interventions, explores common challenges such as issues of causality and contextual variation, and outlines the rationale for the study's chosen evaluation strategy.

2.6.2 Challenges in Evaluation

Although the ideal in educational evaluation is to establish clear causal relationships, achieving this is complex in practice. A central challenge is distinguishing correlation from causation: improved outcomes among intervention participants may result from unobserved confounding factors rather than the intervention itself (Morrison & van der Werf, 2016; Schlotter et al., 2011). In real-world educational settings, it is impossible to control for all confounding variables.

Selection bias poses a significant methodological challenge in this study. Student advisors recommended participation in the MTWs for certain students, but it was ultimately voluntary, meaning that students self-selected into the programme. This introduces the possibility of systematic differences between attendees and non-attendees. For instance, students who volunteered to participate may have been more motivated or more inclined to seek academic support. These unobserved characteristics are associated with improved performance (e.g., Lishchynska et al., 2023; Rylands & Shearman, 2018). As such, these students might have outperformed their peers even in the absence of the intervention, complicating the attribution of observed effects solely to workshop participation.

The results of this evaluation are not readily generalisable. Different institutional cultures, facilitators, peer dynamics, and student backgrounds can impact the effectiveness of an intervention. So, generalising findings across contexts should be done cautiously.

Furthermore, many important outcomes, such as confidence, a sense of belonging, and metacognitive awareness, are affective and difficult to measure using traditional assessments (Crossley, 2020). Although they are crucial for student growth and

academic performance, these “soft skills” are often not given sufficient weight in traditional evaluations. A mixed-methods design is therefore essential to capture both outcomes and mechanisms. Ideally, a long-term longitudinal study would likely disentangle some of the complications.

2.6.3 Methods for Evaluating Interventions

Randomised Controlled Trials (RCTs) are often considered the gold standard for causal inference (Hariton & Locascio, 2018). RCTs reduce the effects of confounding variables by randomly assigning participants to either a treatment or control group. This method allows researchers to make strong causal claims. However, in many educational settings, RCTs are impractical or ethically problematic, especially when they require withholding beneficial support from students who need it.

Quasi-experimental designs offer methodologically sound alternatives to RCTs when randomisation is not feasible or desirable. These approaches aim to establish causal relationships between an intervention and its outcomes while considering any potential biases (Behi & Nolan, 1996). When properly implemented, these designs can yield rigorous research findings comparable to those of RCTs (Panko et al., 2015; Simon, 2025). Standard quasi-experimental methods include Propensity Score Matching (PSM), Regression Discontinuity (RD), and Regression Point Displacement (RPD).

PSM is beneficial when treatment assignment is non-random and based on self-selection, as it creates statistically comparable groups by matching participants on observed covariates (Rosenbaum & Rubin, 1983; Stuart, 2010). This method is most effective when the researcher has observational data and can measure a rich set of covariates that influence selection (Guo & Fraser, 2014). In contrast, RD designs are best suited to contexts where treatment is assigned based on an explicit cutoff or threshold, such as test scores or eligibility criteria (Imbens & Lemieux, 2008). RD offers strong internal validity around the cutoff but has limited generalisability for those further from the threshold. RPD, on the other hand, is suitable for evaluating single-site or small-scale

interventions where only one or a few units receive treatment, with many untreated units serving as comparators (Wyman et al., 2015). While RPD is less commonly used, it can yield credible results when pre- and post-intervention data are available, and treated and control units are reasonably comparable. Given that participation in the MTWs was voluntary and no eligibility threshold was used, PSM was the most appropriate method to minimise selection bias and create a valid comparison group based on observational data.

In this study, students' National Senior Certificate (NSC) and National Benchmark Test (NBT) scores are used as covariates to estimate their propensity scores, which represent the probability of attending the MTWs. Then, treated and untreated students are matched based on their propensity scores, which allows for balanced comparison groups (Rosenbaum & Rubin, 1983). After PSM, treatment effects are estimated. The Average Treatment Effect on the Treated (ATT) estimates the effect of the intervention on those who receive it, which is most relevant for programme evaluation (Shu & Tan, 2018). As Shu and Tan (2018) stated, the ATT answers the question, "How much did persons participating in the programme benefit compared to what they would have experienced without participating in the programme?" (p.1).

2.6.4 Conclusion for Evaluating Educational Interventions

Evaluating educational interventions requires careful consideration of methodological, practical, and ethical constraints. While RCTs offer the strongest causal claims, they are often unfeasible in educational settings. In this study, PSM provides a robust quasi-experimental approach to reduce selection bias and estimate the causal effect of MTW participation. By integrating qualitative insights, the evaluation also captures important non-cognitive outcomes, resulting in a more comprehensive and context-sensitive understanding of the intervention's impact. This dual approach ensures that the evaluation not only measures effectiveness but also informs theory, practice, and policy.

2.7 Gaps in the Literature

The following subsections identify the main gaps and describe how the present study aims to address them.

2.7.1 Limited Focus on Metacognition and Mathematical Thinking

In South Africa, relatively few interventions aim to ease the transition from school to university mathematics and explicitly focus on developing metacognitive and self-regulated learning skills. Most existing programmes emphasise content mastery and, in some cases, collaborative learning, but neglect the intentional cultivation of metacognitive strategies that enable students to plan, monitor, and evaluate their own learning. Given that such skills are central to independent learning and persistence in the face of academic challenges (Cera et al., 2013; Zimmerman & Moylan, 2009), this omission represents a significant gap in the literature. The current study addresses this by embedding structured metacognitive activities into the MTWs, providing students with opportunities for reflection, self-monitoring, and the development of self-regulated learning strategies alongside their mathematical thinking skills.

2.7.2 Lack of Theoretically Grounded Interventions

Although numerous interventions in South Africa aim to address the challenges of transitioning into university mathematics, many lack an explicit articulation of the theoretical frameworks that inform their design (Msomi & Rzyankina, 2024). This limits the ability to generalise findings, replicate effective practices, and build cumulative knowledge in the field. In response, this study adopts a deliberately bounded set of four complementary frameworks[9], Constructivism, the ZPD, APOS Theory, and CLT chosen not to be exhaustive but because they are (a) design-actionable at the level of task and

[9] Again, it should be noted here that these frameworks did not wholly drive the initial design: much of the early pedagogy followed lecturers' intuition. Other lenses are acknowledged and drawn on in later interpretation.

session planning, (b) jointly cover the core mechanisms targeted (e.g., knowledge construction, social mediation, and conceptual development), and (c) align with the intervention aims of higher-order reasoning, metacognitive regulation, and collaborative problem-solving in first-year mathematics.

In some cases, programme evaluations focus narrowly on practical outcomes, such as pass rates, without linking these outcomes to the theoretical mechanisms through which learning is expected to occur. The present study addresses this by integrating the four complementary frameworks into the design of the MTWs. This theoretically grounded approach not only guides the structure and pedagogy of the workshops but also enhances the potential for rigorous evaluation and replication.

2.7.3 Limited Mixed-Methods Evaluations

Despite the growing emphasis on evidence-based practice in education, many intervention studies, both globally and in South Africa, rely predominantly on quantitative indicators, such as test scores or course pass rates. While these measures are important, they often fail to capture the processes through which learning occurs and overlook affective and metacognitive outcomes such as confidence, sense of belonging, or changes in mindset. Mixed methods designs, which combine quantitative and qualitative approaches, offer richer insights into both the measurable impacts of interventions and the lived experiences of participants.

However, Mixed-Methods approaches remain underutilised in mathematics education research, particularly in developing country contexts. For instance, mixed methods account for only 31% of empirical articles in primary mathematics education between 2002 and 2006, showing a 10% decline over this period (Ross & Onwuegbuzie, 2012). Recent reviews echo the value of mixed-methods yet show they remain underused within South African mathematics-education publications (2007–2015) were dominated by small-scale qualitative studies on secondary teaching and learning (Adler et al., 2016), and in a fast-growing subfield (Artificial Intelligence in mathematics education, 2015–

2023) only about 10% of studies used mixed methods (Opesemowo & Adewuyi, 2024). Similarly, a systematic review of nine flagship education journals (2011–2024) found mixed-methods prevalence rose from 0.64% to 2.97%, yet the overall quality[10] of Mixed-Methods Research remained low (7.85/20), with uneven uptake across subdisciplines (Leadership 2.53%; Curriculum and Learning 1.08%) (Gierus et al., 2025).

This study addresses this gap by combining Propensity Score analysis with reflexive thematic analysis, enabling an examination of the statistical effects of the MTWs and the mechanisms, perceptions, and experiences that underlie those effects.

2.8 Conclusion of the Literature Review

This chapter assembled a coherent lens for understanding how the MTWs can influence students' engagement with first-year university mathematics. It argued that improvement is unlikely to arise from any single strand of theory or practice. Instead, gains become plausible when cognitive, metacognitive, sociocultural/identity, and affective perspectives work together: tasks must invite conceptual reorganisation; students must learn to plan, monitor, and evaluate their strategies; participation must confer recognition and legitimate voice; and the climate must lower threat while signalling that reasons, not merely results, count.

Across the cognitive terrain, APOS, ZPD/mediation, and Cognitive Load Theory jointly specify how concepts should be introduced, coordinated, and consolidated. APOS clarifies the intended conceptual trajectory from actions to processes, objects, and schemas. ZPD positions learning at the edge of independent competence, making progress feasible through intentional guidance and support. CLT shapes the presentation and sequencing of tasks so that extraneous demands are pared back, intrinsic complexity is paced, and germane effort is channelled toward schema construction. These lenses

[10] Quality here refers to alignment with Journal of Mixed Methods Research reporting checklist. The checklist, developed for MM manuscripts, outlines 15 core elements (plus 5 additional items for empirical methodological submissions) covering, among others, a clear rationale for mixing, explicit design specification, points of integration, treatment of methodological challenges, and contributions to the mixed-methods field. A higher score indicates greater alignment with the checklist's criteria (Gierus et al., 2025).

are not interchangeable; each targets a specific design issue, and their balancing act, maintaining genuine complexity while preventing overload, and fostering productive struggle without causing discouragement, must be carefully managed.

The chapter also situated learning as participation. Identity is negotiated in interaction; therefore, facilitation practices and peer norms are not peripheral add-ons, but central design elements. Within the South African context, these commitments are not only pedagogical but ethical. Heterogeneous prior preparation, multilingual realities, and uneven schooling histories heighten the risks of peripheral membership and impostor feelings. The design, therefore, treats low-threat climates, clear expectations, and structured collaboration as constitutive of opportunity, not as afterthoughts. Attention to metacognition and self-regulated learning is similarly necessary: without explicit support to plan, monitor, and evaluate their work, many students struggle to transfer emerging insights into durable capability.

CHAPTER 3: WORKSHOP DESIGN AND IMPLEMENTATION

3.1 Introduction

This chapter describes the design and implementation of the Mathematical Thinking Workshops. It outlines the theoretical foundations that support the workshops, the content and organisation of the sessions, the implementation strategy, and the challenges faced during execution. The workshop design specifically addressed challenges (identified in Chapter 2) such as inadequate foundational knowledge, limited metacognitive skills, and difficulty adapting to abstract mathematical thinking. These are the challenges that are directly aligned with the study's aim of exploring how the workshops can improve students' learning experiences and academic performance.

3.2 Theoretical Foundations

The MTWs were established based on established educational theories emphasising active learning, conceptual understanding, and cognitive development. Along with these theoretical foundations, the organising team's intuition and experiences from decades of teaching university mathematics influenced the workshop's design. They drew on their professional expertise and past experiences with students who had encountered challenges in mathematics. The workshops addressed both pedagogical principles and the practical needs of students, as they were grounded in both theory and experience. The next subsections demonstrate how each theoretical framework guided aspects of the MTWs worksheet design.

3.2.1 Constructivism

As discussed in Chapter 2, constructivist learning theory (Thompson, 2020) posits that students learn best when they are actively involved in constructing their knowledge through experience and reflection, not by passively absorbing information (Mugambi,

2018). The workshop's design was primarily based on constructivism, which included planned peer interactions and a strong focus on student-centred active learning.

The workshops featured minimal facilitator-directed presentations and focused instead on problem-solving activities where students engaged with concepts and articulated their reasoning to one another under the guidance of the facilitator and tutors. Students were organised into small groups to collaborate on problems. They were encouraged to use the workshops as a platform to explain their thought processes, learn from different approaches, and gain a deeper understanding through group discussions.

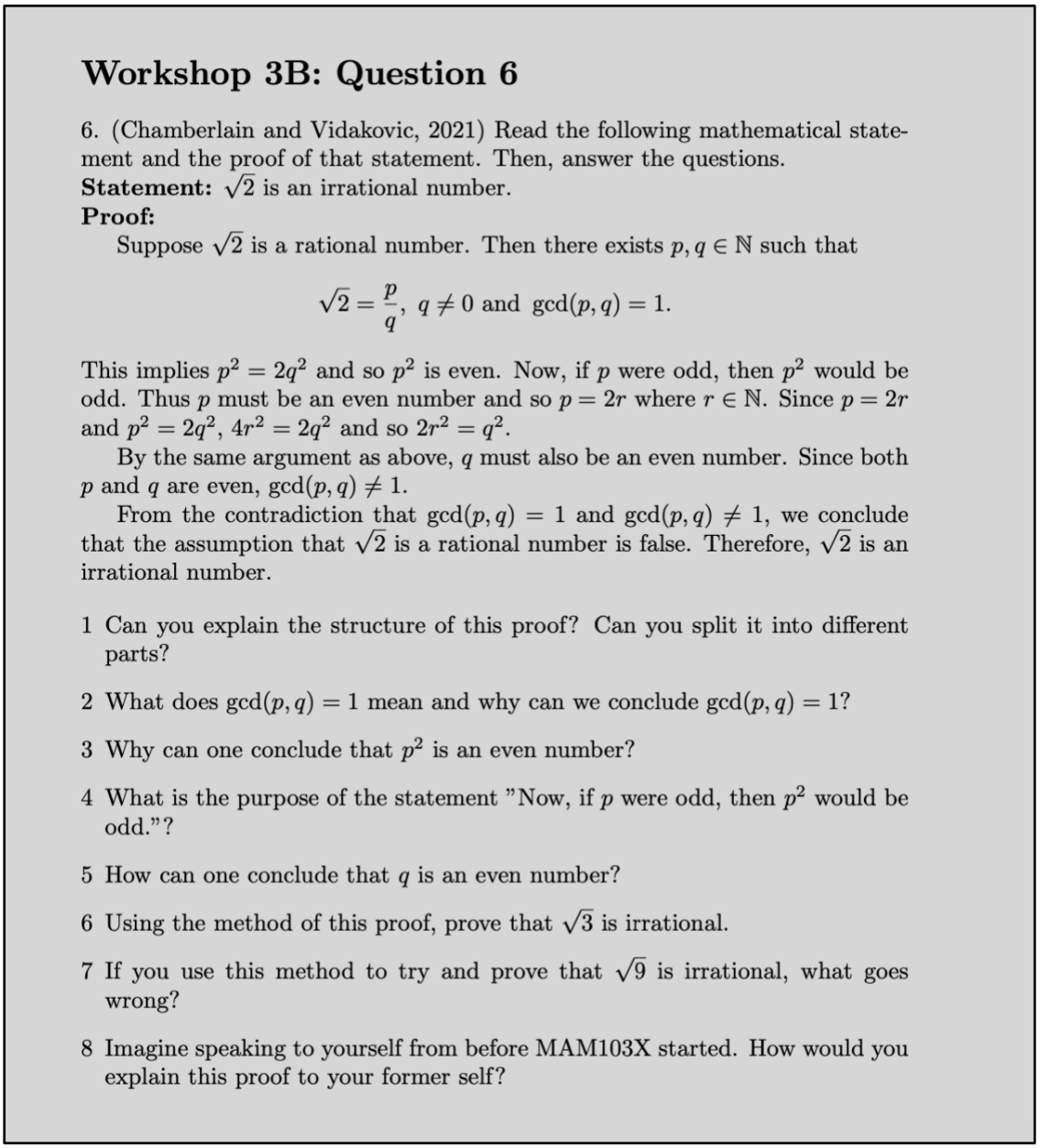
The workshop venue had whiteboards to help students visualise concepts while collaborating with other students. The students used the whiteboards to share their work, clarify their reasoning, and build upon each other's ideas, which created a collaborative learning environment. Different strategies were used to get students to work together. For instance, in one session, the facilitator instructed students to rotate through various roles. These roles included timekeeper, researcher, and seeker, ensuring everyone was involved. The seeker would be sent to examine other groups' work to gather ideas and information that might help their group view things from a different perspective. The whiteboards served as a space for students to engage in mathematics and a platform for applying and critiquing various learning approaches.

The facilitator emphasised the difference between constructing knowledge together as a group and understanding it individually. While reaching a group consensus was necessary, the goal was for each student to develop a personal understanding of the material rather than merely relying on the group's solution. Strategies were implemented to ensure that everyone in the group participated and understood the group's work. Active engagement by every student was essential to ensuring that everyone internalised the concepts discussed. The phrase "individual within the community, rather than individual apart from the community" was often used by the facilitator to show the collaborative yet personal nature of learning mathematics.

3.2.2 APOS theory (design implications and concrete use)

APOS theory, whose theoretical background was discussed in Section 2.5.3, functions as a design heuristic in this section. Worksheet tasks were designed to progress learners from executing steps (Action) to coordinating those steps mentally (Process), to treating the coordinated process as a manipulable entity (Object), and finally to integrating that entity within a broader network of ideas (Schema). In practical terms, this meant scripting worksheet arcs that begin with doable procedures, inserting prompts that require learners to name and operate on the emerging process, and then press for transfer across near and far-variants so that the object is situated within a broader conceptual structure. Facilitation likewise reflected this progression, prompts such as “state the process you just used,” “reverse the direction of what starting action would generate this process,” and “if this process were a single object, what transformations on it are legitimate” served to make the transitions observable and accessible during group discourse.

The sequence of questions asked in Worksheet 3B, as shown in Figure 2, exemplifies this design.



Workshop 3B: Question 6

6. (Chamberlain and Vidakovic, 2021) Read the following mathematical statement and the proof of that statement. Then, answer the questions.
Statement: $\sqrt{2}$ is an irrational number.
Proof:
Suppose $\sqrt{2}$ is a rational number. Then there exists $p, q \in \mathbb{N}$ such that

$$\sqrt{2} = \frac{p}{q}, \ q \neq 0 \text{ and } \gcd(p, q) = 1.$$

This implies $p^2 = 2q^2$ and so p^2 is even. Now, if p were odd, then p^2 would be odd. Thus p must be an even number and so $p = 2r$ where $r \in \mathbb{N}$. Since $p = 2r$ and $p^2 = 2q^2$, $4r^2 = 2q^2$ and so $2r^2 = q^2$.

By the same argument as above, q must also be an even number. Since both p and q are even, $\gcd(p, q) \neq 1$.

From the contradiction that $\gcd(p, q) = 1$ and $\gcd(p, q) \neq 1$, we conclude that the assumption that $\sqrt{2}$ is a rational number is false. Therefore, $\sqrt{2}$ is an irrational number.

1 Can you explain the structure of this proof? Can you split it into different parts?

2 What does $\gcd(p, q) = 1$ mean and why can we conclude $\gcd(p, q) = 1$?

3 Why can one conclude that p^2 is an even number?

4 What is the purpose of the statement "Now, if p were odd, then p^2 would be odd."?

5 How can one conclude that q is an even number?

6 Using the method of this proof, prove that $\sqrt{3}$ is irrational.

7 If you use this method to try and prove that $\sqrt{9}$ is irrational, what goes wrong?

8 Imagine speaking to yourself from before MAM103X started. How would you explain this proof to your former self?

Figure 2: Workshop 3B, Question 6: Proof by Contradiction for the Irrationality (from Chamberlain and Vidakovic, 2021).

In this specific example, the stages of APOS theory are organised as follows:

Action: Students begin by executing step-by-step procedures, such as showing that if 2 is rational, then they can apply parity (even/odd properties) arguments to deduce $p^2=2q^2$ and conclude p must be even. This proof was done in lectures, and students may follow it by relying on memory to follow the procedural steps without necessarily understanding how they fit together logically yet.

Process: Questions such as “Explain the structure of the proof” and “Why is p^2 even?” require students to reflect on the sequence of steps and synthesise them into a coherent logical flow. For example, recognising that assuming 2 is rational leads to a contradiction about *p* and *q* both being even and explaining how parity creates a contradiction. Here, students *interiorise* the proof into a dynamic *Process*, connecting individual steps to the overarching logic of contradiction.

Object: Task 6 (“Prove 3 is irrational”) requires students to *encapsulate* the proof method as a reusable tool. By adapting the same logical structure to a new case (3), they treat proof by contradiction as an *Object,* a static entity that can be manipulated and applied flexibly. To recognise proof by contradiction as a general method, students must abstract away from specific numbers.

Schema: Task 7 (“What goes wrong with 9?”) encourages students to integrate proof by contradiction into a broader logical framework (*Schema*). By testing the method on the number 9 (which is rational), they confront the limits of the technique and refine their understanding of when contradictions arise. This aligns with Schema development, where students coordinate proof by contradiction with other concepts (such as prime factorisation and rationality) to form a cohesive mathematical worldview.

3.2.3 The Zone of Proximal Development (ZPD)

As discussed in Section 2.5.1, Vygotsky's (1978) Zone of Proximal Development (ZPD) complements APOS theory by emphasising the role of external support and scaffolding in learning. In contrast, APOS theory focuses on the internal progression from procedural actions to the formation of abstract schema. These frameworks operate together in the workshops to assist students as they transition from guided problem-solving to independent mathematical reasoning. ZPD focuses on the difference between what a learner understands now and their potential with appropriate support. This theory informed the design of the workshops in several ways, including:

Graded difficulty: The exercises were designed to increase in difficulty gradually. They started with tasks that students could likely complete independently and then gradually introduced more challenging elements to ensure they were working slightly beyond their current understanding, with sufficient support from tutors or peers. For example, Workshop 5, which focused on teaching students how to write proofs, divided a proof by mathematical induction question into three questions of increasing difficulty, as illustrated in Figure 3.

Workshop 5: Question 3

3. An honours class has three students in it: Thando, George and Lizelle. They are not a very organised group, and each class, some of them may not come. There are $2^3 = 8$ different ways that the class could be made up each day:

[Thando, George, Lizelle]

[Thando, George]

[Thando, Lizelle]

[Thando]

[George, Lizelle]

[George]

[Lizelle]

(Nobody shows up)

After a couple of days, a new student, Bongiwe, joins.

1 Without writing down all the possibilities, explain why now there are $16 = 2^4$ possible ways that the class could be made up.

2 Now you are told that some class of k students has 2^k possible ways that it could be made up. Prove from this statement that if there is a class of $k + 1$ students, there are 2^{k+1} different ways that students can attend.

3 With 1 student there are clearly two ways of having the class: They show up, or they don't. Using this fact and what you showed above, prove that for **ANY** number of students n, there are 2^n possible groupings of students in the class.

Figure 3: Proof by induction with varying levels of difficulty.

Scaffolding techniques: The facilitator and tutors strategically used prompts and cues to help students ponder and solve problems without giving away the complete solution. Some of the exercises included worked examples that illustrated how to think through a similar problem. Students could then adapt this approach to the new problem they were working on. See Figure 2 as an example.

3.2.4 Cognitive Load Theory

As discussed in Section 2.5.2, Cognitive Load Theory (Sweller, 1988) emphasises the importance of managing various types of cognitive loads to optimise learning. This theory was applied to guide the design of the workshop exercises, to reduce unnecessary cognitive load and increase relevant cognitive load. Clark and Kimmons (2023) summarised,

> When designing and delivering instruction, educators can reduce extraneous cognitive load by simplifying the presentation of information, engaging in instructional practices that promote germane cognitive load, and adapting instruction to fit learners' zone of proximal development or level of expertise. (p. 2).

Minimising Extraneous Load: The worksheets were designed to minimise elements that might unnecessarily burden students' working memory. This was done by providing (a) clear, concise instructions that did not use jargon unnecessarily; (b) a clean, uncluttered layout with no irrelevant or distracting elements; and (c) a focused approach, with each exercise and workshop targeting a single mathematical concept, where possible, to avoid cognitive overload.

Maximising relevant load: The worksheets focused on (a) building on what students already knew, with exercises that often linked new concepts to students' existing mathematical understanding, which helped to create connections and facilitate schema development; and (b) varied practice, where exercises incorporated different problem

types and representations of the same concept to help learners to develop flexible schemas and apply their knowledge in various situations. The exercise in Figure 3 exemplifies this well.

An example of building on students' prior knowledge can be found in Workshop 1 of Semester 2, where the mathematical thinking skill for the workshop was "making mathematical connections". The mathematical topic for the workshop was the Fundamental Theorem of Calculus, and students were reminded of the concepts they had learnt in the previous semester to minimise the load.

Welcome back!

It's a new semester, a new course and a new season. You will be covering a lot of novel material this semester, but it's important to reflect back on what you learned last semester in order to take stock of the tools and knowledge that you gained and think about how to use it this time around.

In this worksheet we'll be looking at some connections between the two fundamental concepts of calculus: integration and differentiation.

1 First connection:

If $\frac{d}{dx}(F(x)) = f(x)$ then we say that $F(x)$ is an anti-derivative of $f(x)$. We then write

$$\int f(x)dx = F(x) + C.$$

a. Write down an anti-derivative of the function $\frac{1}{x^2} + \frac{1}{\sqrt{x}} + \sqrt{x} + x^2$. Is the anti-derivative unique?

b. Show that $\sqrt{1-x^2} + x \arcsin x$ is an anti-derivative of $\arcsin x$.

c. Verify that $\int \ln x dx = x \ln x - x + C$.

Figure 4: Question linking new concepts to existing mathematical understanding.

This sequencing and layout helped manage cognitive overload and supported schema development. As shown in Figure 4, the worksheet was explicitly designed to link new mathematical concepts with students' existing knowledge. This approach aimed to enhance their conceptual understanding and help manage cognitive load.

3.2.5 Theoretical Foundations Summary

The design of the workshops integrates various educational theories to create a cohesive learning environment. APOS Theory structures the progression of mathematical understanding from procedural to abstract levels. The scaffolding provided by ZPD ensures that students receive the appropriate support as they tackle increasingly complex problems. CLT ensures that instructional materials do not overwhelm students' working memory. These theories come together to form a holistic approach that allows students to build conceptual understanding while receiving customised support incrementally. These theories are also the hypothesised mechanisms linking workshop participation to the outcomes measured in Chapters 5 and 6.

3.3 Workshop Design

3.3.1 Workshop Structure and Format

Students who signed up for the workshops participated in two sessions a week, one on Monday and one on Friday afternoons. Each session lasted between 2 and 3 hours. There are 12 weeks of classes in a semester; therefore, students are expected to go to up to 24 workshops in a semester. These sessions were designed for small groups of 4–6 students to encourage discussion and provide personalised feedback where necessary. To encourage diverse interactions, participants were randomly assigned to groups during the first few workshops. After the students had formed a community, they were allowed to choose their groups during some sessions.

Except for the first three workshops, which were held online due to a protest action on campus, all subsequent workshops were conducted in person at a venue equipped with

several whiteboards to facilitate collaborative work. Each session focused on a specific mathematical thinking skill or metacognitive skill, and students worked on worksheet-based activities tailored to the week’s topic.

Each workshop was different, with no fixed structure adhered to across all sessions. However, certain elements were consistently included, although the order in which they were presented varied depending on the specific objectives of each session. For example, breathing exercises may be incorporated at the beginning, middle, or end of the session, depending on the session's focus. Generally, the workshops aimed to help students focus and encourage innovative thinking. The facilitator often included short videos that provided insights into the day's topic. The sessions also provided time for collaborative work and concluded with time for reflection.

In the first workshop, for example, students were asked to collaboratively compile a list of what they believe are the best practices for effective collaboration and then share their ideas with other groups. To conclude the discussion, they watched a clip of Jo Boaler sharing her insights on collaboration during her guest appearance on Lex Fridman’s podcast.

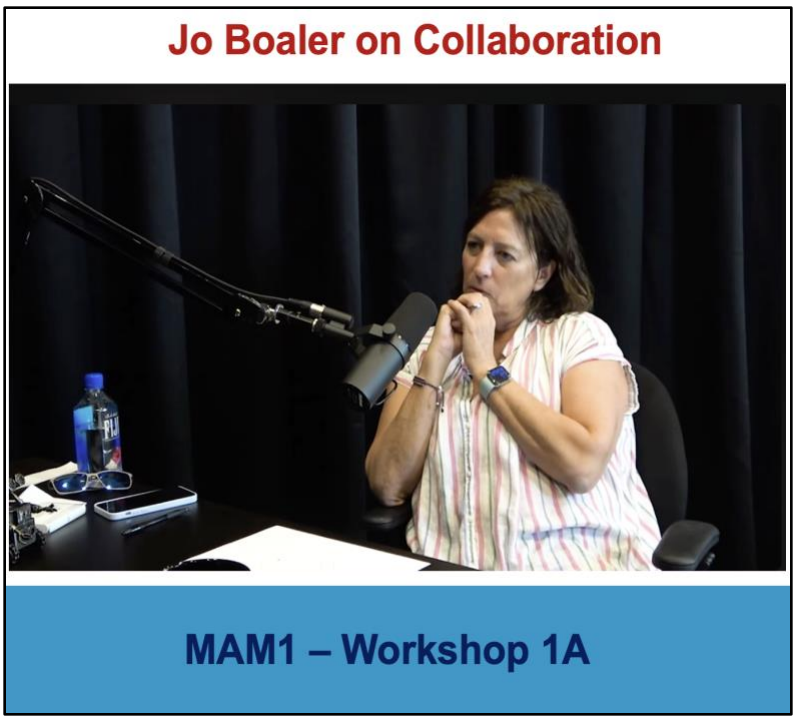


Figure 5: Workshop 1A: Jo Boaler insights on Lex Fridman's podcast

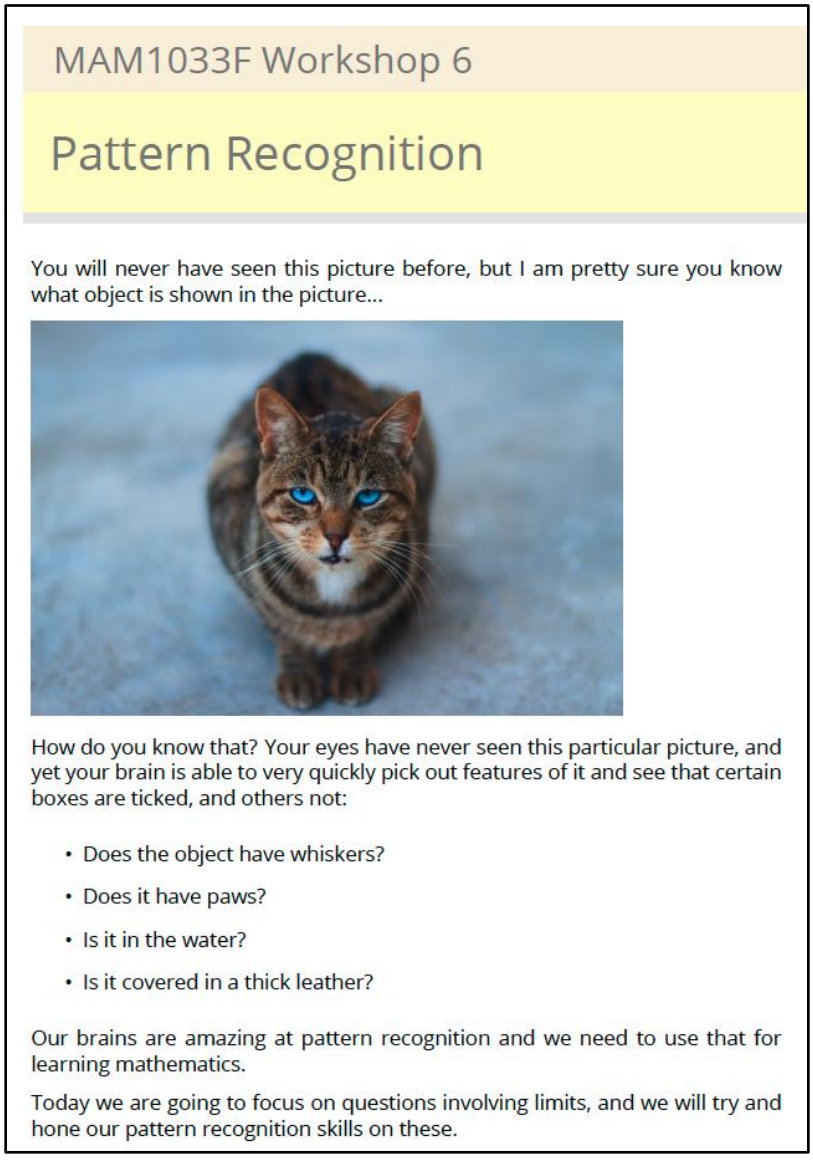
Students collaborated to solve problems and were encouraged to share ideas. The activities varied for each workshop, based on its theme. For instance, when the focus was on effective mathematical writing, groups reviewed each other's work on whiteboards, adding annotations and highlighting areas that needed improvement. These group activities align with constructivist learning theory, which emphasises the importance of collaboration, problem-solving, and knowledge-building through interaction.

Each session typically included a brief warm-in, usually a breathing exercise (2-5 minutes), a short input or video introducing the mathematical thinking practice of the day (delivered by the facilitator or by a mathematician on video, 5 to 10 minutes), collaborative task cycles (60 to 90 minutes), a whole class synthesis (10 to 15 minutes), and an individual reflection or exit prompt (about 5 minutes).

3.3.2 Workshop worksheets and topics

Each session focused on a specific topic related to foundational mathematical thinking or metacognitive skills. The workshop's primary resource was the worksheets, although the

sessions did not necessarily revolve around them. Each worksheet comprised two elements: one covering the mathematical content and the other highlighting the particular mathematical thinking, metacognitive, or study skill through which the content should be viewed. For instance, Worksheet 6 (Figure 6) focused on evaluating limits, and the mathematical thinking skill involved was pattern recognition.



MAM1033F Workshop 6

Pattern Recognition

You will never have seen this picture before, but I am pretty sure you know what object is shown in the picture...

How do you know that? Your eyes have never seen this particular picture, and yet your brain is able to very quickly pick out features of it and see that certain boxes are ticked, and others not:

- Does the object have whiskers?
- Does it have paws?
- Is it in the water?
- Is it covered in a thick leather?

Our brains are amazing at pattern recognition and we need to use that for learning mathematics.

Today we are going to focus on questions involving limits, and we will try and hone our pattern recognition skills on these.

Figure 6: Worksheet 6 focused on pattern recognition in solving limits

As mentioned earlier, the structure of each worksheet included a preamble (Figure 7) that set the stage for the skill and provided motivation for it or featured some external source

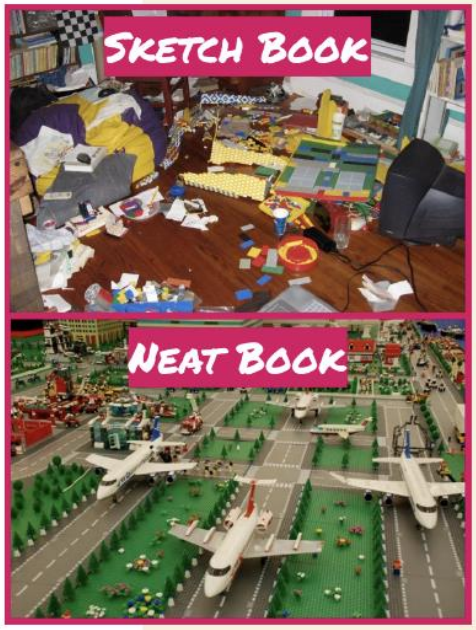
discussing it. Following the preamble, mathematics exercises were designed for small-group collaboration, prompting students to consistently contemplate the skills addressed in the workshop.

MAM1033F Workshop 4A 24 March

Proof by Induction

In this workshop we will be thinking about thinking. There are different ways to get you thinking, and one is to allow you to be free with getting your ideas out.

We discussed before about having two notebooks - a 'sketch book' and a neat book. In the sketch book you want it to be a tangled mess of ideas. Getting ideas out on paper helps to free them from your brain, and you can read them, think further, and move them round as necessary. Staring at a piece of paper, waiting for the ideas to seem perfect in your head is usually not productive. So for this workshop we want you to let those thoughts flow - no matter how useful they seem at the time.



In this first exercise, try to write down whatever comes into your mind. Use words, symbols, and images as you feel necessary. It might be that you know how to solve it and you write down all the steps, or you just have a general sense of what to do, or you aren't sure at all...in which case just right down anything - an example, a trial of a particular number, and see what flows:

1 Prove that for all natural numbers, n:

$$1+2+3+\ldots+n=\frac{1}{2}n(n+1)$$

Figure 7: An example of a preamble, Workshop 4A

In addition to covering core mathematical content, each workshop also addressed a specific aspect of mathematical thinking, study strategies, or metacognitive development. Table 3 provides an overview of the workshops, outlining the mathematical topics and their corresponding metacognitive or mathematical thinking focus. This list is not exhaustive but highlights key themes intentionally incorporated into the design of each session.

Table 3: Overview of Mathematical Thinking Workshops, Mathematical Topics, and Metacognitive Skills

	Mathematical content	**Mathematical thinking or metacognitive skill**
Workshop 1A and 1B	Mathematical statements and definitions of mathematical terms, e.g., what is a theorem, proposition, etc.	Introduction to problem-solving, Understanding mathematical language and notation, Introduction to mathematical proofs
Workshop 2A and 2B	Set notation, intervals, and functions; absolute values and inequalities.	Strategies for reading mathematics and using resources effectively. How to read and write mathematics.
Workshop 3A and 3B	Absolute values, functions; Proof by contradiction	Communicating mathematical ideas effectively; strategies for approaching mathematical problems; how to think like a mathematician.
Workshop 4A	Proof by Induction	Thinking about thinking: having a sketchbook and a neat book.
Workshop 5	Proof Techniques	What it means to prove something, understanding proofs, and proof techniques.
Workshop 6	Evaluating limits	Pattern recognition: working together as a group.
Workshop 7A and 7B	Evaluating derivatives, optimisation problems	Pattern recognition: Discovering for yourself.
Workshop 8A and 8B	Related rates questions: Proving some results related to the Mean Value Theorem	Using the textbook and guidelines to solve problems: how to understand proofs.
Workshop 9	No specific mathematical content	Practices for learning independently.
Other Workshops	Various mathematical contents	Reflecting on mathematical problem-solving strategies, time management and study skills; Using feedback constructively; Making mathematical connections.

3.4 Implementation Strategy

3.4.1 Recruitment of students

Student advisors identified at-risk students using NSC Mathematics scores and encouraged them to enrol in the Mathematical Thinking Workshops. The primary criterion for identifying at-risk students was an NSC Mathematics score below 85%, selected both for its predictive correlation with first-year BSc Science Mathematics performance (Shay et al., 2020). Attending the workshops was voluntary but strongly recommended for identified students. Those who opted in were expected to attend at least 80% of the sessions to remain in the programme, although not all students met this attendance requirement.

3.4.2 Role of Facilitators and Support Staff

The workshops were conducted by one facilitator and supported by three tutors. The tutors supported students' problem-solving by reviewing their work with them and posing prompting questions that encouraged critical thinking and guided them toward discovering solutions, rather than simply providing direct answers.

3.4.3 Feedback and Improvement

Each week, the team responsible for designing the workshops and worksheets met to discuss the previous workshops, plan for the next ones, and refine the worksheets based on what they had learnt from earlier sessions. Some of the topics that were discussed were the difficulty of the questions in the worksheets, the number of questions, the wording of questions, and how questions emphasised particular thinking processes or skills. The team then employed an iterative method to refine the worksheets until they were satisfied with them.

3.5 Challenges

3.5.1 Timetables

Finding two-hour slots within a whole academic week proved challenging, particularly as the students were enrolled in different degree programmes. The organising team worked closely with individual departments to make sure the workshops were accessible and that all students could attend them twice a week. In some instances, the students' curricula were adjusted, allowing them to take specific courses later in their programme. The process was coordinated with the student advisors of each department to ensure that students' programmes would not be extended. In some cases, however, students needed to add an extra year to their studies in order to accommodate the workshops.

3.5.2 Engagement, Motivation, and Adapting to Student Diversity

Maintaining student interest over time was a significant challenge. It was essential to ensure that there was sufficient variety in the workshops to keep students engaged and prevent repetition. Strategies such as varying session activities and introducing new elements, such as interesting facts or physical props, helped sustain interest. While initial enthusiasm was high, keeping students engaged over two 12-week semester courses was not always easy. Motivation also dropped during periods of academic stress, such as midterm exams, leading to irregular attendance. To tackle this issue, the facilitator incorporated new components, such as guest speakers (e.g., a psychologist from the Mathematics and Applied Mathematics department) and interactive activities utilising physical objects (e.g., balloons for related rates problems). Students entered the workshop with diverse backgrounds and different levels of confidence. The workshop's adaptable structure enabled customisation to cater to these diverse needs, but maintaining adequate challenges for each student necessitated constant effort.

3.5.3 Coordination with the Mainstream Curriculum

Aligning workshop content and pace with the formal mathematics curriculum was an ongoing constraint. If topics are out of step with lectures or assessments, students may struggle to see relevance and, in some cases, deprioritise attendance. Coordinating topic sequences and assessment windows with course conveners proved necessary to preserve perceived alignment.

3.5.4 Student Workload and Competing Priorities

Workshops were added to an already heavy first-year load. During peak assessment periods, ungraded workshop activity competed with graded tasks in other courses, and students sometimes chose to focus on imminent deadlines. This pressure may contribute to dips in attendance and reduce cognitive bandwidth for demanding problem-solving.

3.5.5 Drop-in/Drop-out Participation

A portion of students treated the workshops as optional or attended irregularly. Variable attendance disrupted group continuity and weakened rotating roles, which in turn reduced the time available for more in-depth tasks.

3.5.6 Stigma and Perception

Some of the students perceived the workshops as remedial. This framing may discourage them from attending and may have dampened motivation among those who did, particularly higher-attaining students who may have worried about the signal the workshops sent about their ability. Reframing the workshops as “add-ons” that develop higher-order reasoning helped, but the perception remained a persistent challenge in the background.

3.6 Conclusion

This chapter outlines the design and implementation of the Mathematical Thinking Workshops, detailing their pedagogical foundations, session structure, and practical rollout. Grounded in constructivist principles and informed by APOS theory, ZPD, and Cognitive Load Theory, the workshops aimed to support students' transition to university mathematics by promoting conceptual understanding, metacognitive awareness, and collaborative problem-solving.

While worksheets formed a visible component of the sessions, they were often de-emphasised in practice, serving more as prompts for group discussion and exploratory thinking than as exercises to be completed for correctness. The workshops prioritised engagement, flexibility, and peer learning to create a space for students to reflect, strategise, and build confidence in their mathematical reasoning.

Several challenges arose during the implementation process, including timetable clashes, inconsistent attendance, and the need to strike a balance between structure and adaptability. Even so, the workshops provided a valuable and responsive learning environment that aimed to address both cognitive and affective dimensions of learning.

CHAPTER 4: METHODOLOGY

4.1 Introduction

This chapter presents the research methodology employed to evaluate the impact and experiences of Mathematical Thinking Workshops designed to support first-year university students in their transition into university mathematics. A case study mixed-methods approach is adopted, which combines quasi-experimental techniques with qualitative analysis to capture both measurable academic effects and the affective, metacognitive, and social aspects of the intervention.

Section 4.2 describes the research philosophy, paradigmatic perspective, and the role of the researcher. Section 4.3 then explains the research design, which includes both quasi-experimental and qualitative case study components. The subsequent sections describe the research context and participants (Section 4.4), the methods used for data collection (Section 4.5), and the analysis of the data (Section 4.6). Then, in Sections 4.7 and 4.8, the study's validity, trustworthiness, and limitations are discussed. Finally, in Section 4.9, ethical considerations are addressed.

The methodological choices made in this study are informed by the research questions and build on the theoretical and empirical insights discussed in Chapters 1 and 2. For ease of reference, the research questions guiding the analysis are:

Quantitative Research Questions:

1. To what extent do Mathematical Thinking Workshops impact students' academic performance in first-year university mathematics when controlling for pre-university academic achievement?
2. How does the impact of the workshops vary across students with different pre-university mathematics achievement levels?

3. To what extent do the Mathematical Thinking Workshops impact students' performance on higher-order assessment items in first-year university mathematics?

Qualitative Research Questions:

1. How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics?
2. What aspects of the workshops do students find most beneficial for their mathematical development?
3. What challenges and barriers do students experience in applying the strategies taught in the workshops?

4.2 Research Paradigm (Pragmatic Worldview)

Morgan (2007) defines research paradigms as "systems of shared beliefs among a community of scholars" (p. 65). This study is grounded in a pragmatic research philosophy, which emphasises methods and approaches best suited to addressing the research problem rather than strictly adhering to a single paradigm (Brierley, 2017; Tashakkori & Teddlie, 1998). Patton (1990) similarly points out that the pragmatic paradigm demonstrates more of a concern with applications, what works, and solutions to problems. This philosophy contrasts with positivism, which favours quantitative methods and deductive reasoning, and with constructivism, which favours qualitative approaches and inductive reasoning (Dube et al, 2024). The term "worldview", synonymous with paradigm (Patton, 2002), is defined as "a way of thinking about and making sense of the complexities of the real world" (Patton, 2002, p. 69). The pragmatic paradigm is particularly suitable for this study, as it focuses on practical outcomes while acknowledging the objective and subjective nature of educational research.

The pragmatic approach aligns with the study's mixed-methods design in several ways: (a) Ontologically, it accepts singular and multiple realities open to empirical inquiry

(Kaushik & Walsh, 2019) and recognises that, while academic performance can be measured objectively, students' learning experiences are inherently subjective. (b) Epistemologically, it encompasses both objective and subjective knowledge (Creswell & Plano Clark, 2017), enabling the combination of statistical analysis of performance data and interpretive analysis of student experiences. (c) Methodologically, it supports the use of both quantitative and qualitative methods (Creswell, 2015) to provide a comprehensive understanding of the intervention's effectiveness.

Pragmatism enables the integration of quantitative and qualitative approaches, thereby enhancing the strengths and mitigating the weaknesses inherent in each approach (Shannon-Baker, 2015). Hence, this research paradigm enables this study to (a) address selection bias in workshop participation through statistical analysis, (b) use qualitative investigation to capture the complexity of student experiences, and (c) combine both forms of evidence to evaluate the impact of the MTWs. At the same time, Pragmatism is not without limitations. Its emphasis on “what works” can encourage an overly instrumental focus that prioritises practical outcomes over deeper engagement with underlying ontological and epistemological assumptions. In a study such as this, which draws eclectically on sociocultural, cognitive, and information-processing theories, a pragmatic stance may risk glossing over conceptual tensions between these frameworks by subsuming them under a broad commitment to usefulness. Moreover, the flexibility associated with Pragmatism can invite criticism that methodological choices were made on the basis of convenience rather than on more tightly specified theoretical criteria.

4.2.1 Role of the Researcher

As the researcher, I held two roles: I delivered lectures on mathematics at the university and served as a member of the design team for the intervention. Since 2017, I have been instructing within the Department of Mathematics and Applied Mathematics at the University of Cape Town. My academic and professional interest in how students learn comes from my teaching experience and my previous research, which includes a Master’s degree focused on growth mindsets in first-year university mathematics students. These

experiences have shaped my commitment to improving the transitional experience for students entering university-level mathematics.

I was part of the core team that introduced the Mathematical Thinking Workshops evaluated in this study. I co-designed the workshops, developed task sequences and worksheets aligned with the theoretical framework, and tutored during workshop sessions alongside the designated lead facilitator, although I was not the primary facilitator. To minimise power dynamics and social desirability bias in the qualitative component, I did not conduct the focus group interviews myself. Instead, a qualified psychologist external to the academic programme conducted all focus group discussions. I developed the interview questions, prepared the interview guide, and discussed the intent of each question with the external interviewer to ensure alignment with the study's aims while preserving participant comfort. Given my involvement in the workshop design and my academic position, I was aware of the potential for bias, particularly when interpreting the qualitative data. To mitigate this, I engaged in iterative peer debriefing with my three PhD supervisors throughout the data analysis process. This procedure helped ensure that emergent themes were grounded in participants' accounts rather than shaped by my expectations, and that my proximity to the intervention enhanced contextual insight without compromising transparency, ethical rigour, or interpretive integrity.

4.3 Research Design

This study employed a convergent parallel mixed-methods design, utilising a quasi-experimental approach and a qualitative case study to investigate the effects of the MTWs.

4.3.1 Quasi-Experimental Design

Quasi-experimental research designs are valuable tools for estimating the causal impacts of interventions without random assignment (White & Sabarwal, 2014). They are commonly employed in evaluating educational programmes (Gribbons & Herman, 1996).

In the present study, random assignment to the MTWs was not feasible because students identified as at-risk were advised to opt in, and participation was voluntary. Moreover, random assignment would not be appropriate or ethical in this study, as students are recommended to the workshops based on their anticipated needs. This design included two cohorts. Cohort A (treatment group) consisted of all the students who attended additional MTWs, and Cohort B (control group) consisted of all the students who did not participate.

The student advisors recommended the workshops to students identified as at-risk based on their NSC Mathematics scores. These students were then given the choice to attend the additional workshops. However, once they opted to attend, the workshops became compulsory. As a result, students in Cohort A (workshop attendees) are expected to have lower average NSC Mathematics scores compared to non-attendees. To account for this, propensity score methods were used to create a comparison group (Cohort B) with similar NSC Mathematics scores. In addition, overlap-weighted models were fitted to target the Average Treatment Effect for the Overlap (ATO), focusing inference on regions of common support and enhancing balance between the two cohorts. Primary estimands are Average Treatment Effect on the Treated (ATT) for overall assessments and higher-order items. To examine heterogeneity by NSC Mathematics band (70-84 vs 85-100), overlap-weighted models targeting ATO were fitted.

4.3.2 Case-Study Mixed-Methods Approach

The research design for this study integrates case-study and mixed-methods designs as shown in Figure 8. Integrating a case study and mixed methods is well-suited for an in-depth exploration of complex phenomena (Creswell & Plano Clark, 2017). The design was selected for its capacity to provide a comprehensive understanding of the phenomenon (the effects of Mathematical Thinking Workshops on students' learning during the transition to university mathematics) by integrating qualitative and quantitative research methods (Guetterman & Fetters, 2018; Johnson & Onwuegbuzie, 2004).

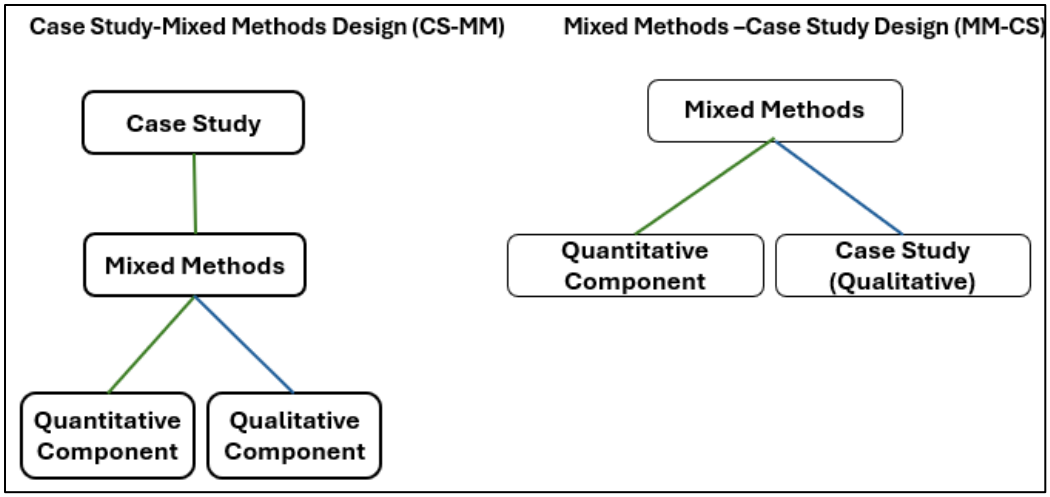


Figure 8: Integrating case study and mixed methods designs (Guetterman & Fetters, 2018).

In this case study, the "case" refers to the Mathematical Thinking Workshops implemented at a specific South African university to support first-year science students transitioning from school to university mathematics. The case study focuses on understanding how these workshops influence students' mathematical thinking, metacognitive skills, academic performance, and overall experiences and perceptions of the intervention.

The following parameters define the case:

- The Intervention: The MTWs, including their design, implementation, and theoretical foundations (i.e., Constructivism, CLT, ZPD, APOS Theory).
- The Context: The University of Cape Town, with its unique challenges such as systemic inequalities, resource disparities, and a diverse student population.
- The Participants: First-year science students who voluntarily participated in the workshops and those who did not (the comparison group).
- The Time Frame: The period during which the workshops are conducted (the 2023 academic year) and evaluated (from 2023 until 2025).

As Yin (2014) articulated, the case study approach allows for a detailed, contextual analysis of a limited number of events or conditions and their interrelationships. This

approach is particularly suitable for this study, as it facilitates an in-depth exploration of the unique context of first-year university MTWs taken by students (participants of this research) transitioning from high school mathematics.

As illustrated in Figure 8 and elaborated by Guetterman and Fetters (2018), there are two common choices for integrating case studies and mixed methods designs. Their distinction is based on prioritising and integrating qualitative and quantitative methods with respect to the case study. The current study employed the Case Study Mixed Methods Design (CS-MM) rather than the Mixed Methods Case Study Design (MM-CS). This choice was made considering that this study prioritises the case study methodological approach, with a mixed-methods approach added to enhance the depth and breadth of the case study.

A case study design enables an in-depth exploration of complex educational environments (Yin, 2018), particularly when the real-life context is integral to understanding an intervention (Creswell & Plano Clark, 2017). In this study, the university setting allows for a naturalistic examination of MTWs and their impact on students' learning during the transition to university mathematics.

Case studies are often employed in the preliminary stages of research to analyse topics that are not yet fully understood, thereby providing rich, qualitative data that can guide subsequent research (Baxter & Jack, 2008). They can also be instrumental in evaluating the effectiveness of a program or intervention within a specific context (Tashakkori & Teddlie, 2010). Yin (2018, p. 55) notes that "case study research has its own place in doing evaluations" and further adds that case studies explain causal links in interventions, describe interventions, illustrate aspects of an evaluation, and illuminate unclear outcomes. In the current study, the case study approach is particularly valuable as it enables an in-depth exploration of the MTWs, offering insights into how they influence students' understanding and application of mathematical concepts while also evaluating their effectiveness in the specific context of first-year science students at a South African university.

Drawing on these strengths, this study employs a case study approach as the overarching framework. The case study approach enables a thorough investigation of the impact of MTWs on students' learning during their transition to university mathematics. The series of MTWs under study was a pilot project, the understanding of which is still in its early stages of development. Assessing the effectiveness of these workshops will enhance the programme's delivery and inform future research.

The Case-Study Mixed-Methods design adopted for this study encompassed two distinct data collection components. The first component, the qualitative aspect, involved conducting focus group interviews to explore participants' experiences, perceptions, and challenges related to the MTWs. These interviews provided a rich source of detailed insights into participants' viewpoints, facilitating the identification of key themes and patterns. In contrast, the second component focused on the quantitative aspect, which involved analysing test and exam scores.

4.3.3 Convergent Parallel Design

Three basic mixed-method designs are widely recognised: convergent parallel, explanatory sequential, and exploratory sequential (Creswell & Plano Clark, 2017). These designs differ in the timing and integration of qualitative and quantitative components. The convergent parallel design collects and analyses both types of data simultaneously, while sequential designs use one method to inform or explain the other (Fetters et al., 2013).

This study considers both quantitative and qualitative methods as equally important in addressing the overarching research question: the quantitative data provided evidence of an association between workshop attendance and academic performance, while the qualitative data captured students’ experiences, perceptions, and the mechanisms behind any observed outcomes. Moreover, data collection for both strands occurred within the same academic year, allowing for concurrent timing. As such, a convergent parallel design was employed to triangulate qualitative and quantitative data. Data from

both strands were collected and analysed separately, then merged at the interpretation stage through systematic comparison and contrast.

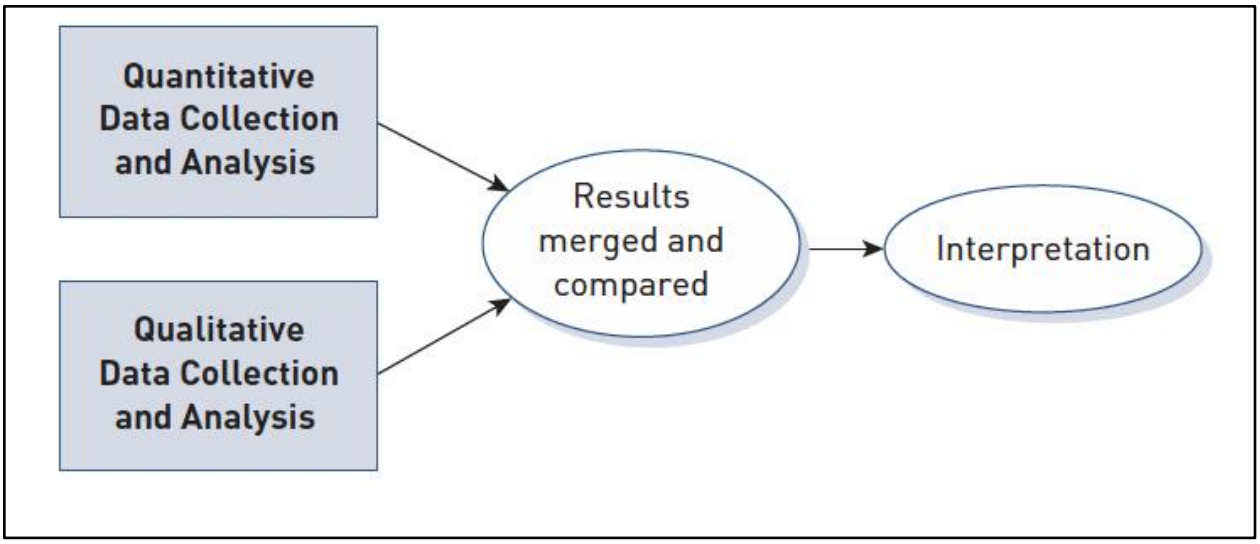


Figure 9: The Convergent Design (Creswell & Plano Clarke, 2017).

This design is best suited when the researcher intends "to obtain different but complementary data on the same topic" (Morse, 1991, p. 122). It also brings together the strengths and weaknesses of quantitative methods (e.g., quantitatively objective measures and qualitatively subjective interpretation) (Patton, 2002). Johnson and Christensen (2008) posit that "the core assumption of this form of inquiry is that the combination of qualitative and quantitative approaches provides a more complete understanding of a research problem than either approach alone" (p. 32).

4.3.4 Case Study Context and Setting

The Mathematical Thinking Workshops are designed for first-year students enrolled in Bachelor of Science majors (e.g., Physics, Astrophysics, Statistics) at the University of Cape Town. These workshops are offered as standalone interventions to support the first-year mathematics course for science students. The course, which enrols approximately 600 learners, attracts learners from diverse schools and backgrounds in South Africa and internationally. This course has been identified as a high-risk course with significant

failure rates, which poses obstacles to student retention and progression (Shay et al., 2020).

The course is structured around two cohorts: Cohort A, comprising students who attend the additional Mathematical Thinking Workshops, and Cohort B, comprising those who do not. At the beginning of the academic year, students with NSC Mathematics scores lower than 85% are recommended to attend the additional workshops. In some cases, depending on the degree program timetable, students may need to drop one of their courses to accommodate the additional workshops.

Aside from attending the workshops, all other academic activities remain the same for both cohorts. Specifically, they attend the same lectures and tutorials and complete the same assessments, including weekly online assignments, two written class tests, and a final exam at the end of each semester. The lectures, delivered in a traditional classroom style, consist of five 45-minute sessions per week, with 200–400 students in each lecture venue. Additionally, all students are required to attend a 1-hour 45-minute tutorial session weekly. These tutorial sessions are held in smaller venues accommodating 40–90 students, with two or three tutors available to assist students in completing tutorial worksheets. While the tutorial sessions are, in theory, compulsory, because no work has to be handed in, the attendance is not vigorously policed, and attendance can vary significantly throughout the semester.

The course content covered throughout the year includes differential and integral calculus, complex numbers, linear algebra, and differential equations. The specific content and structure of the Mathematical Thinking Workshops were discussed in detail in Chapter 3.

4.4 Sampling and Participants

The population for this study consists of first-year science students transitioning from secondary school to university mathematics. The quantitative sampling was based on

students who attended at least 80% of the workshops and were entering university for the first time. Using PSM with NSC Mathematics, NSC English, NBT Mathematics, and NBT Quantitative Literacy scores as covariates, these students were then matched with mainstream students (i.e., those who did not participate in the workshops). Consequently, only students with NSC and NBT data were included in the study's quantitative component. The study's qualitative component only included students who attended the workshops and participated in focus group interviews. Workshop attendance did not imply research participation; consent was obtained separately and the students were informed that participation in research had no impact on grading.

The total population was 538 students, and 61 of them attended the workshops. However, for the quantitative analysis, only students who attended at least 80% of the workshops were included (although attending the workshops was compulsory, not all the students managed to attend at least 80% of the workshops). To provide clarity on who attended the MTWs, both school background and demographic composition of workshop attendees are described below.

Of 61 workshop participants, 3 had attended Quantile 1 schools, 8 Quantile 2 schools, 13 Quantile 3 schools, 9 Quantile 4 schools, and 15 Quantile 5 schools, while the remaining 13 attended private or non-South African schools. This distribution indicates that approximately one-fifth of participants (11 out of 61) came from under-resourced schools (Quantiles 1-2), with additional representation from moderately resourced (Quantile 3) and well-resourced (Quantiles 4-5) public schools, as well as private schools.

UCT also uses a redress-based admissions framework. Redress Category 1 includes applicants with at least one parent classified as Black under apartheid; Category 2 includes applicants with at least one parent classified as Coloured (and the other not Black); and Category 3 includes applicants with at least one parent classified as Indian (and the other not Black or Coloured). Among workshop attendees, approximately 67% of the students who participated in the MTWs were Redress Category 1, 20% Category 2, 3% Category 3, 3% international students, and 7% were admitted under the Open

category. While redress categories do not map perfectly onto school quantiles, they remain strongly associated with historical and structural inequality in access to schooling.

Quasi-experimental designs aim to identify a comparison group that has baseline (pre-intervention) characteristics similar to those of the intervention group (Creswell & Plano Clark, 2017). The comparison group is designed to illustrate what would have occurred if the intervention had not taken place (White & Sabarwal, 2014). Propensity Score Matching (PSM) was employed to establish the comparison group for evaluating the Mathematical Thinking Workshops. PSM is a valuable technique for constructing comparison groups in quasi-experimental program evaluations when authentic experimental designs are not feasible (Gakhar et al., 2010).

According to Randolph et al. (2014), PSM matches treatment cases with control cases based on their propensity scores, which are derived from relevant covariates. PSM is a valuable tool for researchers and practitioners because it strengthens causal arguments in quasi-experimental and observational studies by balancing treatment and control groups across multiple covariates (Connelly et al., 2013; Randolph et al., 2014). There is a risk of self-selection because students are self-selected into the MTWs. However, PSM helps mitigate the effects of this bias by balancing observed covariates and constructing a comparable control group. NSC Mathematics, NSC English, NBT Mathematics, and NBT Quantitative Literacy scores were used as covariates because they are excellent indicators of academic achievement in first-year mathematics at the University of Cape Town (Shay et al., 2020).

Although variables such as gender, school quintile, parental education, and other socio-demographic factors can, in principle, influence student behaviour and performance, these data were not incorporated into the matching procedure for two primary reasons. First, the available administrative dataset did not contain complete information for these variables for the full cohort, which would have resulted in substantial case loss or biased estimation if incorporated into the propensity score model. Second, the decision to prioritise NSC Mathematics, NSC English, NBT Mathematics, and NBT Quantitative

Literacy scores was motivated by empirical evidence demonstrating that these indicators are the strongest and most consistent predictors of first-year mathematics performance at UCT (Shay et al., 2020). Including variables that are either incomplete or weakly predictive risks inflating model variance without improving balance on the characteristics most relevant to mathematical achievement.

To examine variation across NSC Mathematics bands (70–84 vs 85-100), PSM within each band was not applied because the within-band samples were modest; caliper matching would have discarded many cases and produced unstable post-match diagnostics. Instead, overlap-weighted models were fitted using propensity scores estimated from the covariates. Overlap weights focused inference on regions of common support and targeted the ATO. Band-specific overlap-weighted treatment effects and the Treatment × Band interaction was estimated using weighted linear models with heteroskedasticity-robust standard errors, which preserved more information than within-band matching while achieving strong covariate balance within each band.

4.5 Data Collection Methods

To assess the impact of the MTWs on students and their perspectives on the workshops, two complementary strands of data were collected: quantitative and qualitative.

4.5.1 Quantitative Data Collection

The quantitative component aimed to compare the academic performance of students who attended the Mathematical Thinking Workshops with that of students who did not. Pre-university scores, including NSC Mathematics, NSC English, NBT Mathematics, and NBT Quantitative Literacy, were collected from the university's student records and used as covariates in the propensity score analysis to construct comparable treatment and control groups. Students without NSC or NBT data were excluded. Course performance data (comprising two class tests and the final examination for each semester) from the first-year mathematics module served as outcome measures for estimating the effects of

the workshop. Both workshop attendees and non-attendees completed the same class tests and final examinations at the same sittings of the module. Workshop attendance registers defined the treatment cohort, with inclusion restricted to students who attended at least 80% of MTWs sessions.

4.5.2 Qualitative Data Collection

The qualitative data collection aimed to explore students' experiences, perceptions, and the perceived impact of the MTWs. In-depth interviews are a fundamental source of evidence in case study designs (Creswell, 2012; Yin, 2014), as most case studies involve human phenomena. A psychologist who is external to the degree programme conducted semi-structured focus group interviews with students who attended the workshops. These interviews explored how students experienced the workshops, how they perceived changes in their mathematical thinking and problem-solving skills, and what they considered the challenges and benefits of participating in them. The interviews were audio-recorded and transcribed for analysis.

The workshop students were recruited to the focus group interviews via emails from the researcher, who they had met in the workshops, inviting volunteers to share their views on the workshops. Each focus group interview lasted 45 to 60 minutes. The interviews followed a semi-structured format, which offers flexibility and depth while still staying on topic. Researchers often get valuable insights that they did not anticipate from this approach (Adams, 2015). Students were assured that their participation and responses would not affect their academic outcomes.

Six focus group interviews were conducted with 17 unique workshop participants across both semesters; two participants attended two sessions each (one student joined Focus Groups 2 and 3, and another Focus Groups 5 and 6), resulting to 19 total interview attendances. In Week 8 of Semester 1, a three-student group (n = 3) discussed their experiences with the Mathematical Thinking Workshops, including what they found beneficial and challenging, and how participation affected their mathematical thinking, confidence, and collaboration. In Week 10 (n = 2), a second interview revisited these

same lines of inquiry to corroborate and extend the initial insights. By Week 12 (n = 2), the conversation shifted to how group work functioned in practice, the psychological dimensions of learning within the workshops, and whether the sessions were helping to bridge the gap between school and university mathematics.

Semester 2 focus group interviews began in Week 4 with four participants (n=4), focusing on the role of group collaboration in deepening mathematical understanding, comparisons between the workshops and other learning modes, the shift from memorisation to comprehension, and perceived effects on confidence and problem-solving. In Week 5 (n = 3), the discussion examined how interactive group work influenced understanding, alongside the attendant challenges and benefits. The final interview, held in Week 11 with five participants (n = 5), explored students' overall experiences, the extent to which sessions supported their understanding of difficult and complex concepts, and the impact of peer interactions on their learning.

4.6 Data Analysis

Consistent with a convergent parallel design, qualitative and quantitative data analyses were carried out independently.

4.6.1 Quantitative Data Analysis

This section outlines the quantitative strategy used to estimate the effects of MTWs on academic performance. All quantitative analyses were implemented in Python, and the fully annotated code is provided in Appendix H.

Propensity Score Estimation and Propensity Score Matching

To evaluate the impact of the Mathematical Thinking Workshops on students' academic performance, a quasi-experimental approach using Propensity Score Matching (PSM) was employed. PSM is a widely used technique in educational and social science

research for estimating treatment effects in observational studies where random assignment is not feasible (Rosenbaum & Rubin, 1983; Austin, 2011b). It enables the construction of a control group with characteristics comparable to those of the treatment group on observed covariates.

In this study, the propensity score, the probability of a student *i* participating in the workshops given pre-intervention characteristics, was estimated via logistic regression using pre-university achievement covariates (NSC Mathematics, NSC English, NBT Mathematics, and NBT Quantitative Literacy). Formally, let $T_i \in \{0,1\}$ indicate workshop participation and $X_i = (NSC\ Math, NSC\ Eng, NBT\ Math, NBT\ QL)$ pre-university covariates. The propensity score $e(X_i) = Pr(T_i = 1|X_i)$ for student *i* is estimated via logistic regression:

$$logit\big(e(X_i)\big) = \alpha + X_i^T\beta$$

X_i is the covariate vector for student *i*, and β is the logistic-regression coefficients vector estimated from the pre-intervention data (Austin, 2011a; Rosenbaum & Rubin, 1983; Zubizarreta et al., 2023).

The distribution of propensity scores between the treatment and control groups was examined after estimating the propensity scores for all students. This was done to assess the common support (overlap) assumption, which is critical for valid (estimation of) causal inference (Stuart, 2010). Initial inspection revealed a lack of sufficient overlap, especially at the ends of the propensity score distribution (see Section 5.3.1 and Figure 12 for a detailed discussion of this). This indicates that some students who attended the workshop did not have anyone in the control group with similar scores. The lack of overlap presented a threat to internal validity because treatment effects could not be reliably estimated for participants outside the region of common support.

To address this issue, nearest-neighbour matching with a caliper was used. The nearest neighbour pairs each treated student with the control student, whose propensity score is

closest. This type of matching helps to reduce bias due to covariate imbalance (Caliendo & Kopeinig, 2008). A caliper with a width of 0.2 standard deviations of the propensity scores restricts matches to only those within a reasonable range of similarity (Austin, 2009). This technique helped eliminate poor matches that could distort estimates of treatment effect and ensured that matched groups had covariate balance.

After matching, balance diagnostics were conducted to assess the effectiveness of the matching procedure. Standardised mean differences (SMDs) for each covariate were computed before and after matching; this is reported in Chapter 5. Matching was considered successful if the SMDs were reduced to below 0.1, indicating negligible imbalance (Rubin, 2001; Stuart, 2010). The final analysis involved estimating the Average Treatment Effect on the Treated (ATT) by comparing semester test scores and exam results of matched workshop participants and non-participants.

In the context of this study, ATT attempts to measure the effect of the Mathematical Thinking Workshops on the university mathematics performance of students who attended the workshops. Formally (Austin, 2011a; Imbens & Wooldridge, 2009),

$$ATT = E[Y_1 - Y_0 | D = 1]$$

Where:

- Y_1 is the potential outcome under treatment, i.e., the potential test/exam score while attending the workshops.
- Y_0 is the potential outcome without treatment, i.e., the potential test/exam score without attending the workshops.
- D=1 indicates that treatment was received.
- $E[Y_1-Y_0|D=1]$ is the expected outcome difference given that an individual received the treatment (D=1).

In practice, Y_1 and Y_0 are not observed simultaneously for the same individual, so the ATT is estimated using observational data through matching to approximate the

counterfactual condition (Y_0 for the treated group). The matching process aims to find control units similar to the treated units based on observed covariates, thereby approximating $E[Y_0|D=1]$.

The ATT is then estimated as $ATT = \frac{1}{N_{treated}} \sum_{i \in T} \left(Y_{1i} - Y_{0i,matched} \right)$

- $Y_{0i\ matched}$ represents the outcome (class test/exam score) for the control unit matched to the i-th treated unit.
- $N_{treated}$ is the number of treated units.

Variation Across NSC Mathematics sub-groups

Investigating heterogeneity by prior achievement required estimating workshop effects separately for two NSC Mathematics bands (70–84 and 85-100). Propensity score matching was not applied within bands because the within-band samples were modest; nearest-neighbour caliper matching would have discarded many cases, yielding unstable post-match diagnostics and sharply reduced precision. To preserve information while maintaining strong covariate balance, I adopted an overlap-weighting approach that targets the Average Treatment effect for the Overlap (ATO), focusing inference on students whose covariate profiles make either treatment status plausible.

Overlap weights were constructed using propensity scores estimated as before, so that treated students received weight $w_i = 1 - e_i$ and controls received weight $w_i = e_i$. These bounded weights emphasise observations in the region of common support and down-weight observations with extreme propensities, thereby improving balance without discarding units. The causal estimand is the ATO (Li et al., 2017).

$$\tau_{ATO} = \frac{E[(\mathrm{Y}(1) - \mathrm{Y}(0))e(X)(1 - e(X))]}{E[e(X)(1 - e(X))]}$$

Band-specific effects were obtained by fitting weighted linear models for each outcome (class tests, exams, and higher-order item percentages) that included treatment, a band indicator, and their interaction, with the overlap weights applied to all observations. Formally, with $B_i = 1$ for the 85-100 band (and $B_i = 0$ for 70-84), the model

$$Y_i = \alpha + \tau T_i + \gamma B_i + \delta(T_i \times B_i) + \epsilon_i$$

was estimated using heteroskedasticity-robust standard errors under the overlap weights. In this parameterisation, τ represents the overlap-weighted treatment effect for the 70-84 band, while $\tau + \delta$ gives the effect for the 85-100 band; δ tests Treatment x Band moderation. In Chapter 5, I will report simple treatment effects within each band with 95% confidence intervals, and adjust p-values across outcomes within each family using Holm's method, and present weighted balance diagnostics (ASMD) by band to demonstrate that covariate differences were reduced to convectional thresholds after weighting.

Analysis of Performance on Higher-Order Questions

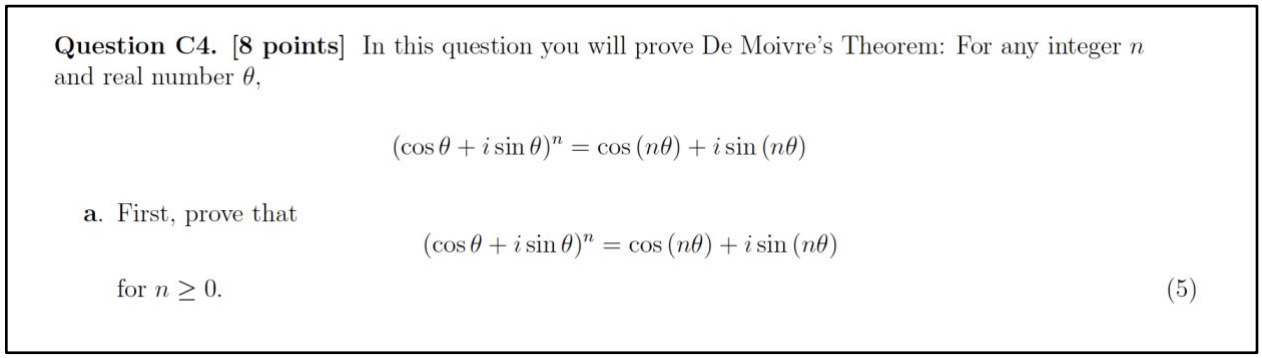
Along with the students' overall evaluation of academic performance, I also conducted a focused analysis to investigate how the Mathematical Thinking Workshops affected their ability to solve higher-order mathematics questions. This analysis aimed to determine whether participating in the workshops was associated with improved performance on tasks that required conceptual reasoning, abstraction, and multi-step problem-solving skills that the workshops were specifically designed to develop.

All the assessments administered during the academic year (Test 1F/S, Test 2F/S, Exam F/S)[11] were uniformly structured with three sections. Section A consisted of multiple-choice questions, Section B included short-answer questions that primarily focused on following procedures and did not require any calculations to be shown, and Section C

[11] Assessment abbreviations: "F" = first semester; "S" = second semester. Each semester includes two class tests (Test 1 and Test 2). "Exam F" and "Exam S" are the end-of-semester exams.

featured long-answer questions that typically assessed reasoning, proof, and conceptual understanding. For this analysis, I focused on Section C of each assessment. However, not all Section C items were included (the complete list of included and excluded items is in Appendix G). I evaluated each question based on the cognitive demand it placed on students. Higher-order questions were those that required students to: (a) construct or analyse mathematical proofs; (b) apply mathematical theorems in unfamiliar contexts; (c) engage in conceptual reasoning; or (d) justify solutions with a mathematical argument.

I excluded questions that, despite being proof-based or conceptually styled, were repeated in lectures, tutorials, or past assessments with slight variation to avoid inflating performances based on recall or rehearsal. For example, De Moivre's Theorem (Figure 10) was covered in lectures and tutorials exactly as it appeared in one of the assessments. A student may be able to reproduce this proof through memorisation alone.



Question C4. **[8 points]** In this question you will prove De Moivre's Theorem: For any integer n and real number θ,

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$$

a. First, prove that

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$$

for $n \geq 0$. (5)

Figure 10:Proof from tutorials and past papers

The exercise in Figure 11 from Test 2F is a good example of a higher-order question because it requires students to apply the Squeeze Theorem in a non-routine context. Students must understand how the graph of cos(1/x) behaves near zero, identify appropriate bounding functions, and justify each step using limit laws to solve the problem.

Question C4. [4 points] Use The Squeeze Theorem to prove that $\lim_{x \to 0} x^4 e^{\cos(1/x)} = 0$.

Figure 11:Higher-order question for Squeeze Theorem

Instead of performing rote calculations, students must reason conceptually about how the function behaves and structure their argument clearly. These are skills that are aligned with the goals of the Mathematical Thinking Workshops. The list of all the excluded and included assessment items is in Appendix G.

Data Preparation for Higher-Order Questions Analysis

Raw scores for the included questions were extracted from the assessment data for each student in the matched sample, based on the Propensity Score Matching (PSM) process described in Section 4.5.1. Only students with complete covariate data and acceptable propensity score overlap in the final matched sample were considered. Each student's scores on the selected higher-order questions were summed by assessment to produce a per-assessment higher-order score.

The analysis proceeded by estimating ATT for each assessment, using the higher-order questions only. Additionally, p-values from independent samples t-tests were reported to assess whether the observed differences were statistically significant. While PSM reduces observed confounding, the results were interpreted as associative, not causal, due to the observational nature of the study and the potential for residual confounding from unobserved variables.

4.6.2 Qualitative Data Analysis

I used reflexive thematic analysis to analyse the transcripts from the focus group interviews. Various researchers have argued that thematic analysis cannot be classified as a method because it provides the basis for many other qualitative analyses (Boyatzis, 1998; Ryan & Bernard, 2000). In fact, the phrase 'thematic analysis' has also been used

to refer to identifying patterns and similarities within qualitative data sets (Boyatzis, 1998; Ryan & Bernard, 2000). Other researchers have argued that thematic analysis is an analysis method (Braun & Clarke, 2006; Nowell et al., 2017).

Thematic analysis is valued for its flexibility, accessibility, and systematic approach to uncovering themes (Kiger & Varpio, 2020). However, Kiger and Varpio (2020) cautioned that more factors should be considered when selecting a method for qualitative analysis. For instance, they stated that the research approach that underlies the analysis should be considered. Braun and Clarke (2006) argued that the researcher's analysis must align with the research paradigm. Several researchers (Teherani et al., 2015; Harper, 2011) corroborated this by stating that the epistemological position should be taken into account, as the philosophical principles underpinning the research will influence the selected qualitative approach. Moreover, researchers might consider their interest in a method and their expertise in using a particular method (Majumdar, 2019). The research questions must also be considered in choosing a method for qualitative analysis.

Braun and Clarke (2006) do not prescribe whether thematic analysis should be done inductively or deductively, as they emphasise flexibility rather than fixed rules. The researcher must select an approach based on the research aims and questions. When thematic analysis is done inductively, themes emerge directly from the data without being guided by pre-existing theories or frameworks. Researchers analyse the data without preconceived notions, allowing patterns and themes to develop naturally based on what they observed. When it is done deductively, the analysis is guided by existing theories, concepts, or predefined categories. Researchers begin with a specific framework or hypothesis and seek data that supports or challenges these predetermined themes. This refers to the fact that the emergence of themes can occur deductively, where the researcher introduces theoretical concepts into the research, or inductively, where themes emerge from the raw data (Joffe, 2012). This flexibility in thematic analysis aligns well with the current study's pragmatic research paradigm, which means "whatever epistemological, theoretical and other commitments are held to facilitate an optimal

engagement with the research aim/questions, it can be incorporated within an analysis" (Dewe & Coyle, 2014, p. 21).

The study employed an inductive approach, which meant that the themes were closely linked to the data and did not attempt to fit into the researcher's pre-existing ideas (Braun & Clarke, 2006). This approach aligns with the view that analysis should not be forced into an existing coding frame (Patton, 1990). This approach aligns with the current study's research aims and questions, as it allows for the capture of students' authentic perceptions and experiences of the Mathematical Thinking Workshops without being constrained by preconceived theoretical frameworks. Braun and Clarke (2006) presented thematic analysis as a method of qualitative analysis, accompanied by a specified set of procedural steps, which are discussed below.

Step 1: The researcher familiarises themself with the data.

I immersed myself in the data (transcripts from focus group interviews) by reading and re-reading them to gain a deep understanding of their content. I engaged in several activities, including transcribing audio recordings verbatim, reviewing the data multiple times, and jotting down my initial thoughts while highlighting interesting or relevant ideas.

Step 2: The researcher generates initial codes.

Step 2 began after I was deeply familiar with the data and had some initial ideas. The task here was to systematically identify and label salient features of the data that are relevant to the research questions (Braun & Clarke, 2006). Data was systematically coded, working through the transcripts line by line or segment by segment. Coding involves identifying and labelling the features of the data that appear meaningful or enjoyable (Stuckey, 2015). I coded using a qualitative analysis software (NVivo) by selecting and highlighting extracts of text and grouping them into a code. I carefully preserved the relevant surrounding texts of the extracts for context. Bryman (2001) warns that context can be lost during the coding process. I applied the codes generously, casting

a wide net, as Braun and Clarke (2006) advised to “code for as many potential themes/patterns as possible (time permitting), you never know what might be interesting later” (p. 96).

Coding was primarily inductive, with codes developed directly from participants’ accounts, while also keeping the research questions in view to ensure relevance. To some degree, the study was also deductive, ensuring that the themes produced informed the research questions. It is common practice to code both deductively and inductively. As Bryne (2022) notes, “Coding and analysis rarely fall cleanly into one of these two approaches, and more often than not, we use a combination of both” (p. 2). Bryne further elaborated on this:

> It is arguably not possible to conduct an exclusively inductive analysis, as the researcher would require some form of criteria to identify whether or not a piece of information is conducive to addressing the research questions and, therefore, worth coding.

In a semantic coding frame, analysis attends to the explicit, surface meanings of participants’ accounts, what is said as stated without inferring beyond the text (Braun & Clarke, 2006). By contrast, latent coding works at a more interpretive level, probing the assumptions and conceptual frames that organise those surface meanings. Braun and Clarke (2006, 2012) caution against saying that codes and themes “emerge”; instead, they are analytic constructions produced by the researcher in dialogue with the data, the research questions, and a theoretical stance. In this analysis, both semantic and latent coding were employed with no attempt to prioritise one over the other.

To complete this step, I coded the interesting features of the data systematically across the entire dataset. The tasks involved tagging parts of the data relevant to the research question with descriptive labels (codes), coding both semantic (explicit) and latent (underlying) meanings, and ensuring that the coding was thorough and consistent.

Step 3: The researcher searches for themes.

Step 3 begins with a lengthy list of initial codes identified across the dataset (Braun & Clarke, 2006). The focus shifted from codes to broader themes. During this phase, I organised into potential overarching themes, combining similar codes to form initial themes that reflected broader patterns or concepts within the data. For this step, I compiled the codes into potential themes and collected all the relevant data for each theme. Activities included organising codes into broader patterns or themes, reviewing how different codes may combine to form overarching themes, and creating thematic maps or tables to visualise relationships between codes and themes.

Step 4: The researcher review themes

Step 4 begins when a set of potential themes has been devised and involves refining those themes (Braun & Clarke, 2006). Patton (1990) suggests that dual criteria should be used for judging potential themes: internal homogeneity and external heterogeneity. Data within themes should cohere together meaningfully, while clear and identifiable distinctions between themes should exist.

Braun and Clarke (2012) proposed a series of key questions that the researcher should address when reviewing potential themes: “Is this a theme (it could just be a code)? If it is a theme, what is the quality of this theme (does it tell me something useful about the dataset and my research question)? What are the boundaries of this theme (what does it include and exclude)? Are there enough (meaningful) data to support this theme (is the theme thin or thick)? Are the data too diverse and wide-ranging (does the theme lack coherence)?” (p. 65)

In this step, I checked the themes to ensure that they are coherent, distinct, and data supported. Activities included evaluating whether themes functioned with coded extracts and the entire dataset, splitting, combining, or discarding themes as necessary, and

ensuring that themes captured the essence of the data and addressed the research questions.

Step 5: The researcher refines and names the themes

Step 5 involved the in-depth refinement of each theme to define its analytical scope and ensure its distinctiveness from other themes. This phase also involves developing concise, engaging, and informative names for each theme that capture their essence and narrative focus. At this phase, each theme and sub-theme is to be expressed in relation to both the dataset and the research questions (Byrne, 2022).

I thoroughly analysed each theme, figuring out what it was about and how important it was. This involved writing a comprehensive description of each theme, explaining how it relates to the research questions, selecting vivid and illustrative data extracts to support it, and assigning a clear and concise name that captures its essence.

Step 6: The researcher produces the report

For this step, I wrote up the qualitative analysis, putting together the thematic narrative with data extracts (Section 6.3). Activities include providing a coherent and logical account of the themes, providing data extracts to illustrate and support the analysis and relating the findings to the research questions and existing literature.

4.6.3 Integration of Quantitative and Qualitative Data

The integration of quantitative and qualitative data in mixed methods research, particularly in case study designs, offers a more comprehensive understanding of complex phenomena (Guetterman & Fetters, 2018). Integration is the hallmark of mixed-methods research (Fetters et al., 2013) and an intentional process by which researchers combine quantitative and qualitative data within a single study (Creswell, 2015). Researchers can achieve integration of mixed-methods research through the

philosophical, design, methods, interpretation, and reporting levels (Moseholm & Fetters, 2017).

At the philosophical level (Section 4.2), this research takes a pragmatic stance, acknowledging the differences between positivism and constructivism while emphasising the benefits of combining them to enhance the strengths and mitigate the weaknesses of each approach (Morgan, 2014; Shannon-Baker, 2015). At both the design level (Section 4.3.3) and the methods level, this study employs a convergent parallel design, where quantitative and qualitative data are collected and analysed separately over a similar timeframe (Fetters et al., 2013). Following the convergent parallel design, both the quantitative and qualitative data are analysed separately, and the results are compared or combined during interpretation. The integration of the data enables the assessment of whether the databases converge, complement, conflict, or diverge (Moseholm & Fetters, 2017).

Integrating results from analyses of separate data components is a common practice in educational research (Bazeley, 2012). Weiss et al. (2005) elaborated:

> This approach can, among other things, enhance validity through data triangulation, delineate overlapping yet distinct aspects of a phenomenon, elaborate on one set of findings using data from another, expand potential findings, and uncover paradoxes and contradictions between results based on different methods. (p. 50)

In discussing strategies for integrating quantitative and qualitative data, Bazeley (2012) emphasised the importance of combining different data types to leverage their complementary strengths and address their non-overlapping weaknesses. He noted:

> The most common form of mixed methods studies similarly involves the use of different data components, comprising separate sources with different types of data or different types of data from within the same source, to complement or

> extend what might be learnt from one or other of those data types independently. These different data are often selected on the understanding that they have complementary strengths and non-overlapping weaknesses, and thus in the expectation that combining them will generate stronger outcomes, that is, outcomes that are better supported by evidence, more generalizable, or both (Johnson & Onweugbuzie, 2004). This strategy for mixing methods might be thoroughly planned, but it also often arises in situations (such as evaluation studies) where the investigator is seeking to gather whatever information might be available from whatever source, as long as it addresses the topic. (pp. 816-817)

Building on this idea, Bazeley (2012) further highlights the growing application of mixed data sources in experimental research and randomised control trials, where the integration of qualitative and quantitative data enhances understanding of both the outcomes and the processes underlying an intervention. As Bazeley explains:

> The practice of employing mixed data sources in evaluation studies is increasingly being applied…. As the course of an intervention is traced and the experience of those involved is assessed, understanding of how the intervention might be contributing to the outcomes can be integrated with knowledge of what that outcome is. Thus, in evaluation studies, analysis of these multiple data sources extends understanding of the processes leading to effectiveness (or otherwise) of an intervention and potentially contributes to explaining causal pathways. (p. 817)

This approach aligns with the current study, where the integration of quantitative data and qualitative data may provide evidence of the MTWs' effectiveness and offer insights into the processes and mechanisms through which the workshops influence students' mathematical thinking, problem-solving skills, and overall academic success.

Creswell and Plano Clark (2017) describe the merging of two databases as a side-by-side comparison. In this approach, the researcher first presents the quantitative statistical results, followed by a discussion of the qualitative findings (e.g., themes) that corroborate

or challenge these results. In this thesis, the quantitative results are reported in Chapter 5, the qualitative results in Chapter 6, and the comparison of the qualitative and quantitative data is made within the discussion and conclusion (Chapter 7).

4.7 Validity and Trustworthiness

Establishing validity and trustworthiness is crucial for ensuring the credibility and rigour of research findings. In mixed-methods research, validity refers to the degree to which the design, data collection, and analysis processes accurately capture the intended constructs and support credible interpretations (Creswell & Plano Clark, 2017). In the quantitative strand, this involves ensuring the internal validity of the results by controlling for confounding variables and ensuring external validity by considering how generalisable the results are. Trustworthiness, on the other hand, pertains to the qualitative strand and includes the credibility, transferability, dependability, and confirmability of the findings (Lincoln & Guba, 1985).

To make quantitative analysis more accurate and less biased, I employed Propensity Score Matching (PSM) to construct treatment and control groups that were similar in terms of observed covariates. This approach reduced selection bias and strengthened internal validity, enabling more reliable causal inferences (Rosenbaum & Rubin, 1983). I also conducted propensity score overlap and balance diagnostics to make sure that the matching procedure is robust and the treatment effect estimates are more credible.

On the qualitative side, I took precautions to ensure trustworthiness throughout the research process. Following Braun and Clarke's (2006) six-phase reflexive thematic analysis, the study's qualitative results became more credible, as codes and themes were grounded in the transcripts of the focus group interviews. To enhance dependability, coding was checked by all three supervisors, with the principal supervisor undertaking a more in-depth review. To reduce social desirability effects and power imbalances, focus group interviews were conducted by an external psychologist with no direct affiliation with the academic programme or intervention.

Triangulation between quantitative and qualitative strands of this study further contributed to the study's methodological integrity. The study successfully captured both measurable outcomes and underlying mechanisms by integrating statistical analysis of academic performance with students' narrative accounts. This combination made the findings more valid and trustworthy (Abowitz & Toole, 2010).

4.8 Limitations of the Research Design

This case study offers valuable insights into the effects of Mathematical Thinking Workshops on students' learning; however, several limitations temper the claims. The findings are context-bound and may not generalise to other universities or student groups. The qualitative strand relied on volunteer participants, potentially over-representing regular attendees or students who particularly enjoyed the workshops. As the qualitative evidence is self-reported, social desirability and recall biases are possible. For instance, some students may have been reluctant to voice criticism for fear of offending the organisers. Using an independent interviewer helped reduce, though not eliminate, potential bias in interviews. Additionally, since members of the research team designed and facilitated the workshops, there is a possibility of researcher expectancy and interpretive bias in both the implementation and analysis.

For the quantitative component, students self-selected into the MTWs, and the analyses were restricted to those who attended at least 80% of the sessions. Consequently, estimated effects may partly reflect pre-existing motivation or engagement among MTW attendees rather than the intervention alone, even after matching to comparison peers. The quantitative samples are relatively small, limiting statistical power and the precision of estimates. Moreover, the outcomes captured here emphasise near-term academic performance and reported experiences; the study does not establish whether observed changes persist or translate into longer-term gains in mathematical thinking or subsequent academic success.

Although propensity-score methods reduce selection bias, they cannot account for unobserved factors (e.g., study habits, time management, resilience) that may influence outcomes. Finally, reliance on formal assessments may understate improvements in informal mathematical practices and shifts in mathematical identity, which are less readily measured.

4.9 Ethical Considerations

Ethical considerations are a set of principles that guide research design and practices, especially in research involving human participants (Ederio et al., 2023). These principles include informed consent, anonymity, confidentiality, and voluntary participation (Bhandari, 2022). This section outlines the steps taken to ensure the study adhered to ethical principles.

4.9.1 Informed Consent

Before participating in this study, all focus group participants provided their informed consent. These participants were provided with a detailed information sheet that outlined the purpose of the research, the procedures involved, and their rights as participants (see Appendix B). They were given sufficient time to review this information and ask questions. Each participant provided written consent, confirming their voluntary agreement to participate.

4.9.2 Confidentiality and Anonymity

Throughout the research process, strict measures were taken to ensure the confidentiality and safety of participants' names and personal information. All data was anonymised by removing or replacing any identifying information. Data was stored securely in password-protected devices and accessible only to the research team. Participants were informed that their responses would be kept confidential and used solely for study purposes. To maintain the confidentiality of individual participants, quantitative results were reported in

aggregate, and pseudonyms were used for qualitative data. Code names were used in the transcriptions of interviews, limiting the identification of data to just one researcher.

4.9.3 Voluntary Participation

It was entirely up to the participants whether to take part in the focus group interviews. They were informed of their right to withdraw at any stage without providing a reason or facing any negative consequences. This clause was clearly communicated during the informed consent process and reiterated at various stages of the research. Participants were also informed that, should they choose to withdraw, their decision would not affect their relationship with the institution or any associated parties.

4.9.4 Ethical Approval

Ethical approval for this study was obtained from the Science Faculty Research Ethics Committee (ethics approval code: FSREC 089 - 2022) at the university before any data was collected (Appendix A). The committee reviewed and approved the research proposal. The approval included the methodology, participant information sheets, consent forms, and data management plans.

4.10 Methodology Conclusion

This chapter outlined the methods used to evaluate the impact of Mathematical Thinking Workshops on first-year Science mathematics students as they transitioned to university mathematics. The study employed a mixed-methods case study design that combined a quasi-experimental approach using propensity score analysis with reflexive thematic analysis of focus group interview transcripts. A pragmatic paradigm guided the integration of quantitative and qualitative approaches. The chapter detailed the research context, the sampling strategies, data collection methods, and analytical procedures applied to both quantitative and qualitative data. Validity and trustworthiness were addressed through triangulation, sensitivity checks to ensure interpretations reflected participants'

perspectives, ethical safeguards, and critical reflexive consideration of the researcher's role.

While limitations such as self-selection and contextual specificity were acknowledged, the design remains well-suited to address the study's research questions and to generate insights into students' transitions to university mathematics. Chapter 3 described the theoretical frameworks that informed the workshop design, the structure and delivery of the sessions, and the implementation strategies adopted to support students to improve their mathematical thinking and metacognitive skills. The following two chapters will evaluate the impacts of the MTWs by analysing student performance data and exploring students' perceptions of their learning experiences.

CHAPTER 5: QUANTITATIVE RESULTS AND ANALYSIS

5.1 Introduction

This chapter presents the study's quantitative findings on the association between participation in the Mathematical Thinking Workshops and first-year mathematics performance. The analysis is presented in four sections. First, Section 5.2 describes the sample and pre-intervention characteristics. Second, Section 5.3 details the propensity score approach used to adjust for observed confounding, including common-support checks, nearest-neighbour matching with replacement under a 0.2 SD logit-caliper, and balance diagnostics based on absolute standardised mean differences (ASMD). Third, Section 5.4 reports Average Treatment Effects on the Treated (ATT) for semester class tests, end-of-semester examinations, and final marks, with cluster-robust inference, Holm control of the familywise error rate within outcome families, and Rosenbaum bounds to quantify sensitivity to hidden bias. Fourth, Section 5.5 examines heterogeneity by NSC Mathematics, explaining the transition from an initial three-band matching strategy to a two-band design (70-84 vs. 85-100) with overlap weighting, which targets the overlap population and achieves an excellent post-weighting balance.

Since the workshops emphasised conceptual reasoning and strategy use, Section 5.6 isolates higher-order (HO) performance by curating Section C assessment items that require multi-step reasoning or proof-adjacent work and constructing an HO% outcome per assessment. The same matched-sample framework is used to estimate HO ATTs, with multiplicity control as above. Section 5.7 then evaluates band-wise HO effects using overlap weighting and a pooled Treatment x Band interaction test.

The intervention was not randomised, and estimates are therefore interpreted as associations conditional on observed covariates. Matching and weighting, as well as balance diagnostics, mitigate bias from observed differences; Rosenbaum bounds provide a calibrated gauge of how strong unmeasured confounding would need to be to overturn significant findings.

5.2 Descriptive Statistics and Sample Construction

The sample for this study comprised 152 first-year Bachelor of Science students at the University of Cape Town. Thirty-one of these students attended the Mathematical Thinking Workshops, while 121 did not. Only students who attended at least 80% of the workshop sessions were included in Cohort A, which was the treatment group for quantitative analysis. This threshold was applied to ensure that the analysis was based on the experiences of students who meaningfully engaged with the intervention and could reasonably be expected to have been exposed to its core components.

The final sample of 152 students was derived from an initial dataset of 532. Several exclusion criteria were applied: only first-time university entrants enrolled in a Bachelor of Science degree were included, resulting in the exclusion of returning students and those registered in commerce or humanities programmes (202 students). BCom students were explicitly excluded because the workshops were not offered to them, owing to timetable clashes within the commerce degree structure. Additionally, students (mostly international students) with missing or incomplete pre-university data, such as NBT or NSC scores, were excluded (104 students), as were records with non-numeric results (e.g., letter grades like ‘A’). While necessary for consistent analysis, these exclusions may introduce some bias. Table 4 summarises the key descriptive statistics for the final treatment group and the control group.

Table 4: Descriptive statistics of Cohort A and Cohort B.

Covariate	**Cohort A[12], those who attended the workshops (n=31)**	**Cohort B, those who did not participate in the workshops (n=121)**	**t-test, p-value**
NSC Mathematics Average Score (Standard Deviation)	84.35 (±6.00)	90.42 (±5.86)	t=5.04, p=7.65e-06
NSC English Average Score (SD)	76.00 (±6.80)	81.00 (±6.98)	t=3.63, p=6.88e-04
NBT Mathematics Average Score (SD)	64.81 (±13.02)	77.78 (±14.90)	t=4.80, p=1.38e-05
NBT Quantitative Literacy Average Score (SD)	55.65 (±19.25)	73.84 (±17.85)	t=4.77, p=2.07e-05

This preliminary analysis suggests that students who attended the workshops had lower pre-university academic scores compared to those who did not. As shown in Table 4, the t-test scores for differences in means exceed 3.60 and are all statistically significant ($p < 0.05$). This finding is noteworthy because historically, NSC and NBT scores have been shown to correlate with university mathematics performance (Shay et al., 2020). These disparities show the importance of accounting for pre-intervention academic achievement when assessing the potential impact of the workshops. The following section reports the results of the Propensity Score Matching (PSM).

[12] Note that some MTWs participants had NSC Mathematics scores above 85% and still chose to attend; participation was voluntary across achievement levels.

5.3 Propensity Score Estimation, Overlap, Matching, and Balance

5.3.1 Propensity Scores Estimation

Propensity scores enable comparisons among students with similar covariates, thereby reducing bias when assessing the impact of the workshops. Rosenbaum and Rubin (1983) defined the propensity score as the probability of getting treatment based on observed pre-treatment covariates. A logistic model was used to estimate the propensity score (Wan, 2025) for attending the Mathematical Thinking Workshops. The model is explained in detail in Section 4.6.1.

Figure 12 shows a density plot used to assess the degree of overlap between treated students (Cohort A) and control students (Cohort B) after estimating the propensity scores. This plot illustrates the distribution of estimated propensity scores for both groups. It also allows for visual inspection of common support, the region where treated and control students have similar probabilities of attending the workshops based on their covariates. Establishing common support is essential to ensure that each treated student can be meaningfully compared to at least one similar control student (control match). If there is not enough common support, it would indicate that some treated students have no comparable peers in the control group, which would make treatment effect estimations less reliable.

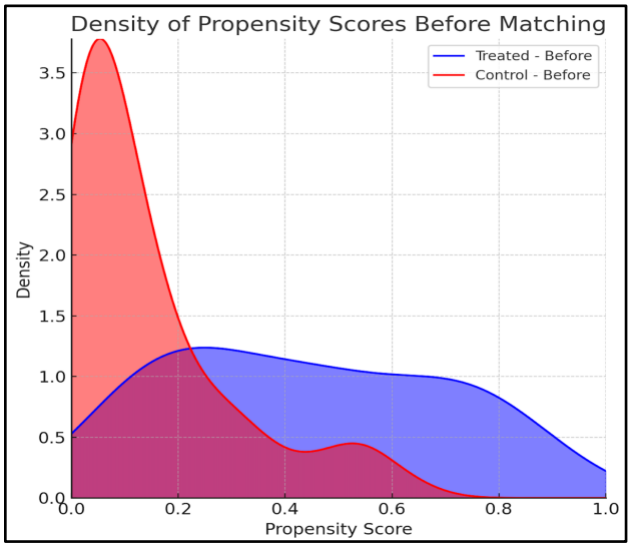


Figure 12: Propensity Scores overlap plot before Propensity Score Matching

While there is a notable difference in the central tendencies of the two groups, there is sufficient (visual) overlap in their distributions to proceed with matching. However, the existence of regions with little to no overlap (particularly towards the right end with propensity scores above 0.8) indicates a potential lack of common support. This may affect the reliability of treatment effect estimates, as there could be inadequate counterparts in the control group for accurate comparison.

There are several approaches to addressing the lack of overlap, including “trimming” (Lee et al., 2011), discarding all units with propensity scores that have either no treated or no control units (Lee, 2005; Rubin, 1977), or implementing caliper matching, where potential matches are dropped if the within-match difference in propensity scores exceeds a certain threshold (Cochran & Rubin, 1973). There are also many other methods (e.g., Dehejia & Wahba, 1999; Heckman et al., 1997; and Ho et al., 2007). This study employed Nearest Neighbour Matching with a caliper, an approach that balances reduction of bias with sample retention, following guidance from Cochran and Rubin (1973). The caliper ensures that only similar pairs are retained, which improves covariate balance and reduces poor-quality matches. The results of this matching are presented in Section 5.3.2.

5.3.2 Propensity Score Matching

To address the first quantitative research question, Nearest-Neighbour Matching (NNM) within a specified caliper width with replacement was used to match each treated subject to a control subject with the closest propensity score. NNM with caliper matches a treated subject to a control subject only if the difference in propensity scores falls within a predefined maximum distance (the caliper). The “with replacement” aspect means a control unit can be matched to more than one treated unit; this increases the chances that all treated units find a match within the caliper.

This effectively addresses the lack of overlap observed in Figure 12, as the caliper excludes units that lack close counterparts in the other group, thereby inherently resolving the issue of overlap. Following Austin’s (2009) recommendation, I set the caliper width as

a fraction,0.2 as recommended by Austin (2011a), of the standard deviation of the propensity scores. The propensity score overlap plots after matching are presented in Figure 13.

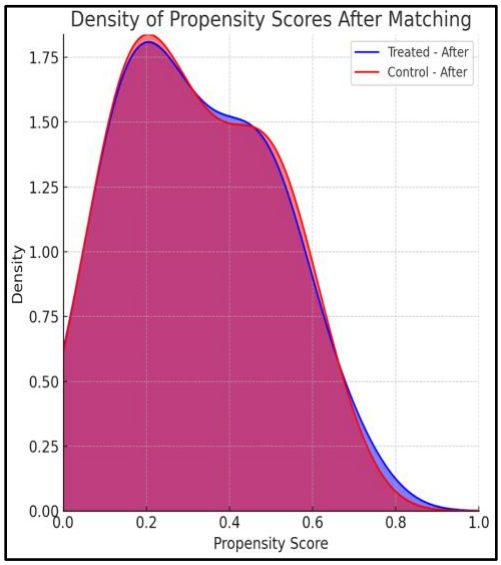


Figure 13: Propensity scores overlap after NNM with caliper

Figure 13 illustrates that the propensity scores of the treatment and control groups are now more similar, which indicates that the matching process has improved the overlap. This shows that the matching procedure has effectively narrowed the difference in propensity scores between the groups, which is crucial for the validity of the quantitative results in this study.

However, good propensity score overlap (a prerequisite for effective matching) does not guarantee the quality of the matching method. The effectiveness of the process depends on how well the two groups are balanced after matching. Balance diagnostics are crucial for verifying that the matching process has mitigated any observed confounding bias and ensured the similarity of covariate distributions between the treatment and control groups. However, they are often improperly conducted and not reported in the literature (Zhang et al., 2019; Ali et al., 2015). Comparing Absolute Standardised Mean Difference (ASMD) is the most common method for assessing balance in covariate distributions after PSM (Zhang et al., 2019; Austin, 2009). This study used SMDs, as they provided a precise,

scale-independent measure of whether students in the workshop and control groups were comparable in terms of the covariates.

Table 5: Absolute Standardised Mean Difference (ASMD) for covariates

Covariate (Average Scores)	Pre-matching treatment Mean (n=31)	Pre-matching Control Mean (n=121)	ASMD (Pre-Matching)	Post-Matching Treatment Mean (n=23)	Post Matching Control Mean (n=17)	ASMD (Post-Matching)
NSC Mathematics	84.35	90.42	1.016	85.61	85.39	0.051
NSC English	76.00	81.00	0.719	77.74	77.96	0.054
NBT Mathematics	64.81	77.78	0.925	67.83	66.00	0.115[13]
NBT Quantitative Literacy	55.65	73.84	1.001	61.13	61.26	0.031

Note: In Table 5, the calculated mean for control group is actually for 24 students (17 distinct students plus the students who were used multiple times)

After matching, the treatment group decreased from 31 to 23 because several attendees lacked comparable controls within the caliper and were trimmed to preserve common support.

[13] Researchers generally pick 0.10 as a threshold to indicate imbalance (Austin, 2009; Chang et al., 2022), but 0.25 is acceptable (Stuart et al., 2013).

Table 6 reports mean covariates and course assessments for three subgroups, trimmed attendees (n=8), matched attendees (n=23), and matched controls (n=17), to show how trimming affected sample composition and performance.

Table 6: Mean covariates and assessment scores after caliper matching.

	Trimmed attendees mean (n=8)	**Matched attendees mean (n=23)**	**Matched controls mean (n=17)**
NSC English	71.00	77.74	78.71
NSC Mathematics	80.75	85.61	84.82
NBT Mathematics	56.12	67.83	64.47
NBT Qualitative Literacy	39.88	61.13	60.59
Test 1F	77.17	83.30	62.04
Test 2F	90.58	89.92	54.08
Exam F	67.25	65.26	63.82
Final F	69.30	68.33	65.70
Test 1S	65.15	54.78	47.02
Test 2S	56.97	53.53	43.99
Exam S	65.62	54.78	50.47
Final S	70.09	60.29	54.06

Note: Unlike in Table 5, the mean scores for the control group (in Table 6) are for the 17 distinct students.

The trimmed attendees (n=8) have weaker pre-university scores than both matched groups (i.e., lower NSC and lower NBT scores), yet they outperformed the matched controls across every assessment. Compared with the matched treated group, they are lower on Test 1F but higher on Test 2F and slightly higher on Exam F and Final F. It follows that dropping these 8 students reduces treated-control gap show in the matched sample (Section 5.4); including them would shift the ATTs upward. Therefore, trimming

in this case yields more conservative ATTs and strengthens the association claims. While propensity score matching is a sensible approach, it may hide very important signals.

I employed the caliper technique for many-to-one nearest neighbour matching (with replacement), enabling the matching of individual control students with multiple treatment students. This approach was chosen to retain as many treatment cases as possible while still enforcing a strict similarity threshold to improve covariate balance between groups.

The balance diagnostics revealed a significant reduction in the Absolute Standardised Mean Differences (ASMD) for all four covariates. Before matching, each covariate had an ASMD value above 0.25; guidelines indicate that ASMDs larger than 0.25 indicate that the groups are too different from one another for reliable comparison (Rubin, 2001). Post-matching, all but one ASMDs fell below the tighter threshold of 0.1 (see Table 5), indicating balance.

With covariate balance achieved between the two groups, the next step is to estimate the Average Treatment Effect on the Treated (ATT). The following section presents the results of this analysis.

5.4 Effect of Workshops on Academic Performance

Interpretive caution

The estimates reported in this section should be interpreted as associations, not causal effects. Although the study employs propensity score matching (PSM) to adjust for observed confounding (e.g., NSC and NBT scores), the non-randomised, voluntary nature of workshop participation introduces the possibility of selection on unobservables (e.g., motivation, time-on-task). Consequently, residual confounding may remain despite good post-matching balance on measured covariates. All inferences below should therefore be read with these limitations in mind.

Estimation strategy

Average Treatment Effects on the Treated (ATT) were computed in Python using the formula introduced in Section 2.5.3 on matched pairs generated by nearest-neighbour PSM with replacement, using a caliper of 0.2 SD on the logit of the estimated propensity score. Of the 31 eligible workshop students, 23 were matched (8 were excluded due to a lack of common support) and paired with 17 unique controls. Post-matching balance met conventional thresholds (Table 5). For inference, I report cluster-robust standard errors to account for control reuse, along with t-statistics, p-values, and 95% confidence intervals. To control the familywise error rate (FWER) across multiple outcomes, Holm's step-down adjustment was applied within outcome families (semester class tests, final exams, and final marks).

Results

Table 7 summarises the ATT estimates for each assessment alongside inferential statistics and standardised effect sizes (Hedges' g and Glass's delta C). The first two class tests of the first semester (Test 1F and Test 2F) show significant, statistically robust gains for workshop participants that remain significant after Holm adjustment: Test 1F (ATT = 21.71; 95% CI [13.69, 29.73]; $p < .001$; pHolm = 1.219 e-04) and Test 2F (ATT = 26.75; 95% CI [14.48, 39.03]; $p<.001$; pHolm = 8.497e-04). The corresponding Hedges' g approximately 1.49-1.81 indicates significant effects.

For the remaining outcomes (final exams and final marks in both semesters, as well as second-semester class tests), point estimates are positive but imprecise and do not survive familywise correction.

Table 7: ATT estimates by outcome with cluster-robust SEs

Assessments[14]	ATT (SE)	t-statistic	p-value	95% CI	Hedges' g	Glass's delta C	p (Holm)
Test 1F	21.71 (3.78)	5.74	3.047e-05	[13.69, 29.73]	1.81	2.44	1.219e-04
Test 2F	26.75 (5.79)	4.62	2.832e-04	[14.48, 39.03]	1.49	1.69	8.497e-04
Exam F	0.91 (4.99)	0.18	0.857	[-9.69, 11.52]	0.06	0.06	0.8575
Final F	2.2 (3.76)	0.58	0.567	[-5.78, 10.18]	0.18	0.21	0.5892
Test 1S	6.39 (5.07)	1.26	0.226	[-4.36, 17.05]	0.4	0.54	0.2270
Test 2S	9.26 (5.53)	1.67	0.114	[-2.46, 20.99]	0.57	0.7	0.2270
Exam S	4.91 (5.38)	0.91	0.375	[-6.50, 16.33]	0.28	0.35	0.7500
Final S	4.93 (4.55)	1.08	0.295	[-4.71, 14.57]	0.34	0.44	0.5892

Note: ATT units are percentage-point differences; effect sizes are reported as Hedges' g and Glass' delta C.

[14] "Final F" and "Final S" are the end-of-semester final marks for each semester, combining the two class tests and the exam. The final mark is 0.3×Class Record + 0.7× Exam or 0.4×Class Record+0.6×Exam whichever is higher. The class record is the average of two semester class tests and WebAssign score, WebAssign are online weekly homework exercises.

Sensitivity to hidden bias

I assessed robustness to unobserved confounding using Rosenbaum bounds (Wilcoxon signed-rank) on the matched pairs (one-sided, testing for improvement among workshop attendees). The tipping point Γ* is the smallest factor by which an unmeasured trait would need to tilt the within-pair odds of workshop participation to make the observed effect non-significant at α=0.05, conditional on the observed covariates.

Table 8:Rosenbaum Sensitivity (Wilcoxon Signed-Rank) on matched pairs

	S (Wilcoxon T^+)	**p at Γ=1**	**Γ* at p=0.05**
Test 1F	265.5	5.3e-05	4.20
Test 2F	237.0	1.67e-04	3.35
Exam F	150.5	0.3518	1.00
Final F	160.0	0.2517	1.00
Test 1S	176.0	0.1239	1.00
Test 2S	193.0	0.0154	1.30
Exam S	171.0	0.1577	1.00
Final S	175.0	0.1302	1.00

Using the 23 matched pairs (drawn from 17 unique controls), the early first-semester class tests are notably robust. For Test 1F, Γ* = 4.20, meaning that even if an unmeasured factor made one student in each pair about 4.2 times more likely to attend the workshop, the advantage would remain statistically credible. For Test 2F, Γ* = 3.35, indicating that the finding remains convincing under a roughly 3.35-fold worst-case tilt in selection odds.

Second-semester Test 2S shows a positive signal but is fragile (Γ*≈1.30). Outcomes that are not significant at Γ=1 (no hidden bias) have Γ*=1.00 by definition (exams and final marks). Together with the ATT results, this pattern supports significant, robust gains on Test 1F and Test 2F, with small or imprecise effects on the other assessments.

Workshop participation is associated with significant improvements on early (first two class tests of the year) formative assessments; effects are robust to moderate levels of hidden bias and to multiplicity control. Effects on exams and overall final marks are small and statistically insignificant.

5.5 Variation Across Student Sub-Groups

This section attempts to answer the quantitative research question:

How does the impact of the workshops vary across students with different pre-university mathematics achievement levels?

The initial plan was to estimate band-specific average effects on the treated (ATTs) using within-band PSM. However, post-match balance diagnostics were unsatisfactory in the original band split (70-84; 85-89; 90-100): several post-match ASMDs exceeded the conventional thresholds for acceptable balance (0.10-0.25). Trimming was not an option, as band-wise samples were already tiny. Because credible causal interpretation requires good covariate balance (Stuart, 2010), I did not use those three-band PSM ATTs. To improve common support and precision, I collapsed NSC Mathematics scores into two bands: 70-84 and 85-100. This collapse improved overlap and enabled a design that prioritises balance over pairwise matching. Rather than matching, I adopted a balancing-weights approach specifically overlap weighting, which targets the region of common support and avoids extreme weights (Li et al., 2017; Austin & Stuart, 2015).

Overlap weighting achieved an excellent balance across all four pre-treatment covariates, with post-weighting ASMDs less than 0.10 (see Table 9).

Table 9: Covariate balance (ASMD) before and after overlap weighting.

Covariate	ASMD (pre-weighting)	ASMD (Weighted)
NSC Mathematics	1.023	0.0
NSC English	0.725	2.33e-15
NBT Mathematics	0.927	1.01e-15
NBT Quantitative Literacy	0.980	7.76e-16

Table 10 reports band-specific overlap-weighted effects for all assessments (estimate, SE, Holm adjusted p), alongside the pooled Treatment × Band interaction p-value. Very small p-values are reported in scientific notation.

Table 10: Workshop effects by pre-university mathematics achievement

	Band 70-84 **Effects [SE] (pHolm)**	**Band 85-100** **Effects [SE](pHolm)**	**p (interaction)**
Test 1F	22.23 [3.437] (pHolm = 1.97e-10)	19.19 [5.05] (pHolm = 1.46e-04)	0.618
Test 2F	31.63 [6.96] (pHolm = 5.50e − 06)	25.42 [5.42] (pHolm = 5.38e − 06)	0.482
Exam F	2.66 [5.54] (pHolm = 1.00)	-1.86 [4.32] (pHolm = 1.00)	0.521
Final F	3.33 [4.66] (pHolm = 0.95)	-0.57 [3.36] (pHolm = 0.95)	0.498
Test 1S	0.59 [8.05] (pHolm = 0.94)	7.66 [4.59] (pHolm = 0.19)	0.455
Test 2S	5.63 [5.61] (pHolm = 0.63)	4.02 [6.29] (pHolm = 0.63)	0.849
Exam S	5.64 [8.12] (pHolm = 0.97)	-2.96 [6.10] (pHolm = 0.97)	0.397
Final S	4.56 [6.69] (pHolm = 0.99)	-0.27 [5.10] (pHolm = 0.99)	0.567

Notes. Effects are simple treatment effects within each band; pHolm adjusts the two-band tests per outcome. The interaction column reports p for Treatment × Band (70-84 vs. 85-100). Estimates are obtained using Weighted Least Squares with robust SEs.

First-semester tests (Test 1F and Test 2F) show large, positive, and statistically significant workshop effects in both bands. Test 1F**:** +22.23 (70–84, pHolm<.001); +19.19 (85-100, pHolm<.001), and Test 2F**:** +31.63 (70–84, pHolm<.001); +25.42 (85-100, pHolm<.001).

For both outcomes, the Treatment × Band interaction is non-significant ($p > 0.48$), indicating no evidence of moderation by prior achievement. Second-semester class tests show small, imprecise positive estimates in both bands, not significant based on pHolm. Exams and final marks estimates are near zero and not statistically significant on either band.

For the outcomes where the workshops demonstrably help (first two class tests of Semester 1), the impacts are positive for both achievement bands and do not differ reliably in magnitude. This suggests that students benefit similarly, regardless of whether they enter with a score of 70-84 or 85-100 in Grade 12 Mathematics. Therefore, the workshop impacts do not vary by pre-university mathematics achievement when analysis is conducted on a well-balanced overlap-weighted sample.

5.6 Performance on Higher-Order Questions

This section presents a targeted analysis of student performance on selected long-answer questions that align with the goals of the Mathematical Thinking Workshops. These questions required higher-order thinking, such as conceptual reasoning, abstraction, mathematical proof, and multi-step problem-solving.

Each assessment in the course was structured into three sections: Section A comprised multiple-choice questions; Section B contained short-answer questions that required brief computations; and Section C consisted of long-answer questions that typically involved

multi-step reasoning, proofs, or conceptual problem-solving. A review of all Section C items across the six assessments (see Appendix G) revealed that while many questions appeared to require proof or conceptual insight, several were drawn almost directly from lectures or tutorials and thus may reflect memorisation rather than independent reasoning. Such questions were excluded to ensure the analysis accurately reflected the kinds of mathematical thinking the workshops aimed to develop. This is discussed in detail in Section 4.6.1, and all the Section C assessment items are presented in Appendix G together with the list of included and excluded assessment items.

For each assessment, I computed a higher-order (HO) percentage score, capping values at 100 to avoid inflation from bonus mark items. The causal estimand is the ATT of workshop participation on HO%, evaluated separately for each of the six assessments. I reused the propensity scores and matching from Section 5.3.

Table 11: Average treatment effects on the treated (ATT) for higher-order (HO) items

	ATT (SE clustered)	**t-statistic**	**p-value**	**95% CI**	**Hedges g**	**Glass delta c**	**p-adjusted holm**
Test 1F	2.536 (6.774)	0.374	0.712	-11.51, 16.59	0.1091	0.1115	1.000
Test 2F	13.527 (5.992)	2.258	0.034	1.10, 25.95	0.7440	0.8065	0.205
Exam F	2.341 (5.893)	0.397	0.695	-9.88, 14.56	0.1301	0.1320	1.000
Test 1S	7.880 (7.737)	1.019	0.319	-8.164, 23.93	0.3294	0.3505	1.000
Test 2S	8.184 (5.699)	1.436	0.165	-3.64, 20.00	0.3539	0.3226	0.825
Exam S	4.537 (5.976)	0.759	0.456	-7.86, 16.93	0.219	0.2454	1.000

Table 11 reports ATTs in percentage-point difference in HO%. 95% Confidence Intervals and p-values come from cluster-robust t-tests. Effect sizes are Hedges' g (pooled SD)

and Glass's delta (control SD). Family-wise error was controlled across the six tests using Holm-Benferroni.

Unadjusted tests indicate positive and statistically significant gains in higher-order items for Test 2F in the first semester (≈ +13.5 percentage points (pp); g ≈ 0.7440). After controlling for six comparisons with the Holm procedure, however, the result no longer meets the α = 0.05 threshold (adjusted p = 0.205). In other words, the estimate remains positive and of moderate magnitude, but the evidence is suggestive rather than confirmatory and is statistically compatible with smaller or even no true effect. Other assessments show smaller, non-significant point estimates.

5.7 Higher-Order Questions (Variation Across Sub-Groups)

To examine whether workshop effects (on HO items) differ by prior achievement, I re-estimated treatment effects within NSC Mathematics bands of 70-84 and 85-100 using overlap weights and propensity scores from Section 5.3.2.

Table 12: Higher-order (HO) performance by prior achievement

	Effect (70-84) [SE]	**pHolm (70-84)**	**Effect (85-100) [SE]**	**pHolm (85-100)**	**Interaction p**
Test 1F	11.26 [7.77]	0.294	3.00 [8.50]	0.724	0.473
Test 2F	19.03 [5.69]	0.002	7.54 [5.06]	0.136	0.131
Exam F	3.74 [6.22]	1.000	-1.09 [5.44]	1.000	0.559
Test 1S	4.05 [8.001]	0.613	12.32 [6.49]	0.116	0.422
Test 2S	6.58[8.18]	0.843	5.52 [6.92]	0.843	0.922
Exam S	-0.23 [9.34]	1.000	0.43 [6.18]	1.000	0.953

Note: Effects are treated - control differences in HO% using overlap weights within each band. pHolm adjusts the two-band tests within each outcome. Interaction p is the approximate test of equality of band effects (Treatment × Band: 70-84 vs 85-100).

Within the 70-84 band, effects are positive across most assessments, with a notably large and statistically strong impact on Test 2F: +19.03 pp (SE=5.69, pHolm=0.002). Effects are directionally positive but not statistically significant (after Holm) for Test 1F: +11.26 pp (SE=7.77, pHolm = 0.294), Test 1S: +4.05 pp (SE = 8.001, pHolm = 0.613), Test 2S: +6.58 pp (SE = 8.18, pHolm = 0.843), and Exam F: +3.74 pp (SE = 6.22, pHolm = 1.000). The Exam S effect is essentially zero (−0.23 pp, SE = 9.34, pHolm = 1.000). These patterns suggest that students with solid (but not top-end) school mathematics benefited most on mid-semester, higher-order tasks, especially in the second term of the first semester.

Within the 85-100 band, effects are generally smaller and less precise. The largest estimates are +12.32 pp on Test 1S (SE = 6.49, pHolm = 0.116) and +7.54 pp on Test 2F (SE = 5.06, pHolm = 0.136). Other assessments show small, non-significant differences: Test 2S +5.52 pp (SE = 6.92, pHolm = 0.843), Exam S +0.43 pp (SE = 6.18, pHolm = 1.000), Test 1F +3.00 pp (SE = 8.50, pHolm = 0.724), and Exam F −1.09 pp (SE = 5.44, pHolm = 1.000). Overall, high-achieving entrants show at most modest gains.

The Treatment × Band interaction tests do not reach conventional significance for any outcome (interaction p ranges from 0.131 to 0.953), with the sharpest contrast again on Test 2F (Interaction p=0.131), indicating a suggestive (but not definitive) difference favouring the 70–84 band over 85-100. For the remaining assessments, interaction p-values (0.422–0.953) imply no reliable differential effect by band.

In this cohort, overlap-weighted estimates indicate that workshop benefits on higher-order performance concentrate among students with NSC Mathematics in the 70–84 range, particularly on second-term test work (≈ +19 pp on Test 2 F after multiplicity adjustment). For 85-100 students, estimated gains are smaller and imprecise. Importantly, exam-level differences remain small and non-significant in both bands. Given modest band-specific treated sample sizes (≈15–16 per band) and correction for multiple testing, these findings

should be viewed as robust directional evidence of band-specific mid-semester gains, rather than uniform improvements across summative assessments.

5.8 Summary of Quantitative Findings

This chapter used a matched-sample design and overlap-weighted analyses to estimate the impact of the Mathematical Thinking Workshops on multiple outcomes across two semesters. Propensity score matching (1:1, with replacement; caliper = 0.2 SD on logit PS) reduced the treatment sample, retaining 23 attendees paired with 17 unique controls and excluding several unmatched attendees to preserve common support. Inference used cluster-robust standard errors and Holm adjustment within outcome families. Because trimming alters the target population and the distribution of baseline characteristics, these exclusions may have materially affected the statistical significance of estimates; the direction and magnitude of any distortion are unknown.

On the matched sample, two formative tests in Semester 1 showed significant and statistically robust gains for workshop participants after Holm correction (Test 1F and Test 2F), while exams and final marks did not differ significantly between groups. Effect sizes for Test 1F and Test 2F were large (Hedges g ≈ 1.5–1.8).

Sensitivity analyses using Rosenbaum bounds on matched pairs indicated that these formative gains (Test 1F and Test 2F) are reasonably robust to hidden bias (critical $\Gamma \approx 4.20$ for Test 1F and ≈ 3.35 for Test 2F). In contrast, outcomes that were non-significant at $\Gamma = 1$ remain so by definition.

To examine heterogeneity by prior achievement, Section 5.5 re-specified the banding to two NSC Mathematics bands and moved from within-band matching to overlap weighting, which achieved excellent covariate balance (all post-weighting ASMD < 0.010). Overlap-weighted models showed consistent improvements on the early formative tests (Test 1F

and Test 2F) across both bands, and the Treatment × Band interaction was not significant (p-interaction>0.48), suggesting no reliable moderation by prior achievement; exams and final marks remained near zero in both bands.

For higher-order (HO) items, the chapter defined inclusion/exclusion criteria to isolate multi-step reasoning tasks, computed capped HO% per assessment, and estimated ATT with cluster-robust inference. Unadjusted estimates showed ≈ +8 to +13 percentage-point gains on Test 2 in both semesters with moderate to large effect sizes; after Holm adjustment, these HO gains were marginal (adjusted p= 0.825 for Test 2S; = 0.205 for Test 2F), and other assessments were not significant.

Overall, the workshops are associated with meaningful improvements on early formative assessments (Test 1F and Test 2F), no reliable differences on final examinations and overall marks, no detectable moderation by prior achievement, and positive but only marginally significant improvements on targeted higher-order items once multiplicity is controlled. Given the observational design and potential residual confounding, as well as the post-matching exclusions noted above, which may have substantially skewed statistical significance, the findings should be interpreted with caution and understood as pertaining to the matched subset rather than all attendees.

CHAPTER 6: QUALITATIVE FINDINGS AND ANALYSIS

Disclaimer on Researcher Role and Reflexivity

The qualitative analysis presented in this chapter is based on the principles of reflexive thematic analysis (RTA), which recognises the active and interpretive role of the researcher in generating knowledge. This approach asserts that researchers construct themes through their ongoing engagement with and interpretation of the data, in contrast to grounded theory, which posits that themes naturally "emerge" from the data (Glaser, 1992; Glaser & Strauss, 1967). As outlined in the later extensions of RTA (Braun & Clarke, 2012, 2019, 2021), reflexivity (i.e., the practice of critically reflecting on one's own assumptions, biases, and experiences) is considered a valuable analytic resource. Rather than being considered a threat to objectivity, the researcher's subjectivity is embraced as an integral part of the analytic process. Therefore, the findings presented in this chapter represent one possible and situated interpretation, co-constructed through sustained engagement with the data.

Given my dual role in the broader MTW project and my non-involvement in data collection, I maintained a reflexive journal after each analytic session (memos on assumptions, early hunches, tensions), used 'critical-friend' debriefs with supervisors to contest early theme shapes, and deliberately sought disconfirming data during theme review. I treat the themes as one theoretically informed interpretation, not the truth of the data.

6.1 Introduction

This chapter presents the qualitative analysis from six focus group interviews conducted across an academic year with students who participated in the Mathematical Thinking Workshops. It aims to explore how students experienced the workshops and how these experiences influenced their learning and engagement with university mathematics. The findings are presented thematically and are grounded in Braun and Clarke's (2006) six-

phase approach to thematic analysis, which includes familiarisation with data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report.

The analysis aimed to address the three qualitative research questions:

1. How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics?
2. What aspects of the workshops do students identify as most beneficial for their mathematical development?
3. What challenges and barriers do students experience in engaging with the workshop content and implementing the learnt strategies?

Through this process, five key themes developed from the data and are presented in the sections that follow. Each section includes subthemes where applicable, with illustrative participant quotes that bring the themes to life and support the narrative. The analysis focuses on what students reported as meaningful in their learning journeys, how the workshop structure and facilitation affected the engagement, and what barriers persisted despite the interventions.

6.2 Overview of Thematic Analysis Process

The analysis presented in this section was conducted using Braun and Clarke's (2006, 2021) six-phase framework for thematic analysis, which is discussed in detail in Section 4.6.2.

1. Familiarisation with the data: transcripts were read multiple times to identify patterns.
2. Generating Initial Codes: Key phrases and recurring ideas were labelled with codes.

3. Searching for Themes: Codes were grouped into overarching themes.
4. Reviewing Themes: Themes were refined and checked for coherence with the data.
5. Defining and Naming Themes: Themes were described and linked to research questions.
6. Writing the Report: Findings were synthesised into a structured narrative.

NVivo software was used to manage and analyse qualitative data systematically. The thematic analysis approach was particularly suitable for this study, as it allowed for an in-depth exploration of students' subjective experiences in the Mathematical Thinking Workshops while remaining grounded in the data. Thematic analysis was employed to identify commonalities and differences in students' perspectives and to generate themes that directly address the study's three qualitative research questions.

6.2.1 Familiarisation with the Data

The qualitative dataset consisted of transcripts from six focus group discussions, each with two to six participants, as described in Section 4.5.2. Only students who had attended the MTWs were invited to participate in these discussions. The focus groups included a diverse range of students enrolled in various science majors, including first-time university entrants and returning students. All audio recordings were transcribed verbatim, and an initial read-through was conducted to familiarise the researcher with the breadth and depth of the dataset. During this process, notes were taken on early impressions, repeated ideas, and potential patterns that could be explored further.

Braun and Clarke (2006, p. 92) noted that familiarisation is not merely about reading but about actively generating meaning and engaging with the data. They described this process as "searching for meanings" and posited that this analysis stage "provides the bedrock for the rest of the analysis." After data familiarisation, initial codes were generated; this step is elaborated on in the next section.

6.2.2 Generating Initial Codes

The initial coding process produced 162 coded text segments and 18 descriptive codes. Some of the most frequently occurring codes are listed in Table 13; the complete list of codes and extracts is provided in Appendix E. Some extracts belonged to multiple initial codes; for example, the extract “The workshops help take away that mindset that you need to cram something to do well” was under both the initial codes *Metacognitive Awareness and Strategic Thinking*, as well as *Deep Conceptual Understanding and Unlearning Rote Habits.*

Table 13: Initial codes generated from the focus group interview transcripts.

Initial code	Number of extracts[15] and an illustrative extract
Collaboration and Peer Learning	28 extracts “It helps you to collaborate with other people, and different people have different perspectives on a specific problem, so it helps you to learn more and grow as a student.”
Metacognitive Awareness and Strategic Thinking	26 extracts “The workshops help take away that mindset that you need to cram something to do well.”
Deep Conceptual Understanding and Unlearning Rote Habits	23 extracts "It's really incredible because we pinned down those misconceptions that are lying there very deeply in our understanding. And we build the foundation very solidly.”
Psychological Safety and Personal Growth	22 extracts “I call it math therapy. It feels like therapy.”

[15] Following RTA, salience is judged interpretively rather than by frequency. Counts of text extracts are descriptive of the dataset handling, not indicators of prevalence or ‘weight’ of a code.

Intentional Design and Inclusive Facilitation	21 extracts “I think the most helpful part for me is when we go as a group to the boards to solve the questions. It really helps because you get to discuss with the other students, and then you get new perspectives on the problem. That's what's helpful for me.”
A Sense of Community, Belonging, and Peer Solidarity	20 extracts “But in the workshop, it's like everyone is so involved in helping one another. If you don't understand something, nobody's judging you when you ask a question. And you can see the amount of effort that's put into preparing the workshops and stuff; it really makes you feel like there are people supporting you as opposed to before we had the workshops.”
Confidence and Resilience in Mathematical Problem-Solving	18 extracts “I just pause and think about the problem before I worry about the solution. So that has helped me to approach things with more confidence and be less intimidated because I've ended up believing that I can solve a problem if I have enough time to think about it and solve it.”
Frustrations, Gaps, and Environmental Barriers	16 extracts "But then you also get to work with people, and the difficulty of it is that you're working with people, and some people won't really participate. Somebody would just tell you they are just tired, and that could be a bit frustrating. It's hard to work with other people because I haven't really worked with anyone in my life. It's either I’m working alone or I’m just

	teaching myself all this stuff. So it's a good experience to work as a group, but then it's quite challenging if you're not used to it.”

Any text that could be useful in addressing the research questions was coded. Data extracts were often assigned multiple overlapping codes to reflect their layered meaning. For example, a student discussing how they learnt to pause and think critically before solving a problem was coded as demonstrating both metacognitive awareness and strategic thinking, as well as confidence and resilience in mathematical problem-solving. This aligns with Braun and Clarke's (2006) sentiment that an extract “may be uncoded, coded once, or coded many times, as relevant” (p. 97).

The codes in Table 13 result from repeated iterations and further familiarisation. Codes that helped answer the research questions were kept, and those that did not were discarded. Thematic analysis is a recursive process, and a researcher rarely follows a linear path through the six phases (Braun & Clarke, 2014); the original transcripts are regularly consulted even after initial codes have been generated.

Once all relevant data items were coded, Phase 3 commenced. This phase is detailed in Section 4.6.2 and the next section.

6.2.3 Searching for and Reviewing Themes

In Table 14, codes were assembled into initial candidate themes. As Braun and Clarke (2006) point out, “a theme captures something important about the data in relation to a research question and represents some level of patterned response or meeting within the data set” (p. 10).

Table 14: Nine preliminary themes and their codes.

Preliminary Themes	**Combined Initial Codes**
Theme 1: Growth in Communication and Mathematical Language	Collaboration and Peer Learning; Confidence and Resilience in Mathematical Problem-Solving; A Sense of Community, Belonging, and Peer Solidarity
Theme 2: Psychological Support and Emotional Safety	Psychological Safety and Personal Growth: A Sense of Community, Belonging, and Peer Solidarity
Theme 3: Alignment with Lectures and Assessments	Intentional Design and Inclusive Facilitation
Theme 4: Frustrations, Gaps, and Environmental Barriers	Frustrations, Gaps, and Environmental Barriers
Theme 5: Metacognitive Awareness and Strategic Thinking	Metacognitive Awareness and Strategic Thinking; Deep Conceptual Understanding and Unlearning Rote Habits
Theme 6: Reignited Interest and Motivation in Mathematics	Psychological Safety, and Personal Growth; A Sense of Community, Belonging, and Peer Solidarity
Theme 7: Confidence in Problem-Solving	Confidence and Resilience in Mathematical Problem-Solving; Psychological Safety and Personal Growth
Theme 8: Facilitator Influence and Inspirational Guidance	Intentional Design and Inclusive Facilitation; Psychological Safety, and Personal Growth
Theme 9: Desire for Continuity and Guided Independence	Frustrations, Gaps, and Environmental Barriers; Metacognitive Awareness and Strategic Thinking

Most codes were associated with a single theme, although some were linked to multiple themes. For example, A Sense of Community*, Belonging, and Peer Solidarity* fell under Preliminary Themes 1, 2, and 6, and *Confidence and Resilience in Mathematical Problem-Solving* also fell under multiple preliminary themes.

The code *Frustrations, Gaps, and Environmental Barriers* became a preliminary theme. It is not unusual for a code to represent an overarching narrative within the data and be promoted as a subtheme or even a theme (Braun & Clarke, 2012). This theme directly informs the third qualitative research question (i.e., what challenges and barriers do students experience in engaging with the workshop content and implementing the learnt strategies?). I merged the code *Metacognitive Awareness and Strategic Thinking* with the code *Deep Conceptual Understanding and Unlearning Rote Habits*; these two codes then became a potential theme, *Metacognitive Awareness and Strategic Thinking*. This theme directly informs the first two qualitative research questions:

1. How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics?
2. What aspects of the workshops do students identify as most beneficial for their mathematical development?

6.2.4 Reviewing Potential Themes

I reviewed the potential themes at two levels. At the first level, I reviewed the preliminary themes for internal homogeneity to ensure coherence within each theme and external heterogeneity to ensure distinction between themes. This level involved examining whether the coded segments accurately represented the essence of each theme and whether each theme consistently captured the underlying meaning across different focus groups. This review led to refining theme boundaries and rewording theme names, where necessary, to better reflect their analytical focus.

For example, the theme *Growth in Communication and Mathematical Language* combines codes related to both interpersonal interactions (such as peer explanations) and formal written reasoning (such as proof clarity). From the beginning, these aspects were grouped under a single theme due to their shared focus on expression and articulation. As the analysis progressed, their connection became even more evident; both verbal and written forms of communication were central to how students expressed and developed their mathematical thinking. Therefore, I retained this theme and refined its boundaries to emphasise expressive clarity and confidence-building in both spoken and written mathematical contexts. Consequently, I renamed the preliminary theme *Growth in Mathematical Communication and Problem-Solving Confidence.* This refined title captured the dual emphasis on students' developing ability to communicate mathematical reasoning and their growing confidence in contributing to collaborative problem-solving environments.

Table 15: Refinement and grouping of preliminary themes into five final themes.

Final Theme	**Combined Preliminary Themes**
Theme 1: Developing Reflective and Strategic Mathematical Thinking	Metacognitive Awareness and Strategic Thinking
Theme 2: Growth in Mathematical Communication and Problem-Solving Confidence	Growth in Communication and Mathematical Language; Confidence in Problem-Solving
Theme 3: Reconnecting with Mathematics Through Alignment with Coursework	Alignment with Lectures and Assessments; Reignited Interest and Motivation in Mathematics
Theme 4: Facilitator Influence and Supportive Learning Environment	Facilitator Influence and Inspirational Guidance; Psychological Support and Emotional Safety
Theme 5: Student Voices on Workshop Hurdles	Frustrations, Gaps, and Environmental Barriers; Environmental Discomfort and Venue Constraints; Desire for Continuity and Guided Independence

While both *Psychological Support and Emotional Safety* and *Reignited Interest and Motivation in Mathematics* addressed affective aspects of students' experiences, I distributed them across two distinct final themes. I merged *Psychological Support and Emotional Safety* with *Facilitator Influence and Inspirational Guidance* to form the theme *Facilitator Influence and Supportive Learning Environment*, which emphasised the role of the facilitator in creating an emotionally affirming and motivational learning space. In contrast, I combined *Reignited Interest and Motivation in Mathematics* with *Alignment with Lectures and Assessments* to form the theme *Reconnecting with Mathematics Through Alignment with Coursework*, to capture how well-timed, relevant workshop content contributed to students' academic engagement and renewed enjoyment of mathematics.

I transformed the preliminary theme, Metacognitive Awareness and Strategic Thinking, into a final theme and broadened it by renaming it *Developing Reflective and Strategic Mathematical Thinking*.

The theme *Student Voices on Workshop Hurdles* was developed by combining three preliminary themes, as shown in Table 15. Although I initially coded these themes separately to reflect cognitive, physical, and future-facing concerns, I found them to converge around students' experiences of obstacles that hindered full participation in the workshops, as well as their proposed solutions for enhancing the learning environment these elements together under one theme allowed for a more integrated narrative that captured both the barriers encountered and the constructive feedback offered.

The second level of review began with the five candidate themes I developed during the first-level refinement. At this stage, I did not generate any new themes, nor did I modify or remove any. Instead, this stage functioned as a validation process, which involved re-reading the entire dataset to assess whether the five themes accurately captured the breadth and depth of participant experiences. The goal here is to ensure that the themes are not only reflected in the specific coded extracts but also resonate with the entire data corpus and that no significant patterns have been overlooked (Braun & Clarke, 2006).

This step also allowed for refining illustrative quotes and identifying additional data segments that supported existing themes. As described by Braun and Clarke (2006), this stage ensures that “the themes work in relation to the data set” and that researchers consider whether “to code any additional data within themes that has been missed in earlier coding stages” (p. 98).

6.2.5 Defining and Naming Themes

As Braun and Clarke (2006) described this phase, ‘define and refine’ means identifying the essence of what each theme is about and determining which aspect of the data each theme captures. Table 16 provides a summary of the five themes, including their definitions and their alignment with the corresponding research questions.

Table 16: Summary of the final themes and alignment with the research questions.

Final Theme	Definition	Research Question Addressed
Theme 1: Developing Reflective and Strategic Mathematical Thinking	Describes students' shift from rote, procedural approaches to more analytical, purposeful, and reflective mathematical problem-solving.	RQ1: How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics? RQ2: What aspects of the workshops do students identify as most beneficial for their development?
Theme 2: Growth in Mathematical Communication and Problem-Solving Confidence	This theme captures how students became more confident in expressing, explaining, and critiquing mathematical ideas, both verbally and in writing, especially in peer settings.	RQ1: How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics? RQ2: What aspects of the workshops do students identify as most beneficial for their development?
Theme 3: Reconnecting with Mathematics	Reflects how the workshops reinforced lecture content, supported assessment readiness, and re-engaged students emotionally with mathematics.	RQ2: What aspects of the workshops do students identify as most beneficial for their development?

Through Alignment with Coursework		
Theme 4: Facilitator Influence and Supportive Learning Environment	Highlights the role of the facilitator in creating a non-judgemental, emotionally safe space where students felt motivated, supported, and seen.	RQ1: How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics? RQ2: What aspects of the workshops do students identify as most beneficial for their development?
Theme 5: Student Voices on Workshop Hurdles	Represents the challenges students faced, such as environmental discomfort, time constraints, and a lack of continuity, and their suggestions for improving the workshops.	RQ3: What challenges and barriers do students experience in engaging with the content and applying strategies?

These final themes passed the “two-sentence” test, which Braun and Clarke (2006) describe: “It is important that by the end of this phase you can clearly define what your themes are and what they are not. One test for this is to see whether you can describe the scope and content of each theme in a couple of sentences.” (p. 92)

6.3 Emergent Themes (Producing the Report)

This section presents the five overarching themes that were developed through the Braun and Clark (2006) RTA:

(1) Developing Reflective and Strategic Mathematical Thinking,
(2) Growth in Mathematical Communication and Problem-Solving Confidence,
(3) Reconnecting with Mathematics Through Alignment with Coursework,
(4) Facilitator Influence and Supportive Learning Environment, and
(5) Student Voices on Workshop Hurdles.

Each theme is detailed below, with supporting codes and illustrative participant quotes.

6.3.1 Theme 1: Developing Reflective and Strategic Mathematical Thinking

This theme was developed through the clustering of initial codes related to metacognitive awareness, strategic thinking, and conceptual breakthroughs. Across focus groups, participants described a deepened awareness of how they were learning mathematics. Rather than simply seeking answers, students described learning to pause, think through the structure of problems, and consciously choose appropriate strategies. This shift from rote procedures to purposeful engagement with the material reflects the development of reflective problem-solving behaviours.

“Palesa” (FG1) illustrated this transition well when she shared,

> I used to have this mindset that when I wanted to solve a problem, I just used trial and error. I rushed to do that, but now there’s a lot of pause and thinking. There’s

> a lot of, OK, interpret this problem in a different way so that it appears to me simpler and easier to think about.

At the same time, it is important not to equate all “trial and error” with poor practice. Research on mathematical problem-solving treats the conjecture-test-revise heuristic as legitimate (Singh et al., 2018). What expert problem solvers do is structure their trials, generate hypotheses, and use feedback from partial results to refine their plan (Pólya, 2014; Schoenfeld, 1985, 1992, 2012). In other words, the pedagogical goal is not to abandon experimentation, but to transition from unguided/random trial-and-error to hypothesis-driven exploration, integrated with planning and monitoring. Such productive struggle involving attempts, evaluations, and adjustments supports schema building and conceptual insight when it is purposefully facilitated and cognitively scaffolded (Hiebert & Grouws, 2007; Kapur, 2014).

“Juan” (FG1) corroborated this, noting, “I just pause and think about the problem before I worry about the solution. So that has helped me to approach things with more confidence and be less intimidated.” Students also described using the group work in the workshops as metacognitive monitoring, a way to check methods, uncover misconceptions, and calibrate what they knew, as “Jane”(FG3) noted, “Working as a group has its benefits, you get to realise your mistakes, your misconceptions, and what you know versus what you should have known.”

These moments of insight, described by students as turning points, were often tied to the workshop structure, which prioritised conceptual understanding over procedural repetition. As “Musa” (FG1) explained further,

> They give you a large variety of problems and they teach us how to think. We are learning how to learn, which is something you don’t really get anywhere in the university. We don’t really get that in tutorials; we don’t really get that in lectures.

His reflection highlights the distinctive epistemological environment created by the workshops, one that prioritises learning processes rather than content alone.

This is corroborated by “Amanda” (FG4), who reflected that:

> Our understanding started shifting when we had to literally analyse line by line of a proof…we had to know how this line shifted from line one to line 2, the *how* was important, that’s how we shifted from memorising to understanding.

“Walter” (FG2) added:

> They’ll bring us to the more useful systematic way of approaching problems, which you first need to understand the nature of the problem itself so that you are able to do what is necessary to solve the problem, not just take out that procedure that you have crammed in your brain and attempt.

A key feature of this theme was the move from external validation (e.g., correct answers) to internal reasoning and problem interpretation. “Aphiwe” (FG 5) shared a vivid example related to integral calculus:

> We were given shapes and told to figure out the volume… We hadn’t even been taught that method yet, but we used what we understood to solve it. It wasn’t about memorising; it was about making sense of it.”

This example illustrates how students internally developed problem-solving strategies and transferred them to unfamiliar contexts.

Musa added, “We have a lot of explanations on how to recognise patterns to solve problems before we actually get to solve problems. Which is something we don’t really get from the lectures and stuff.”

These accounts demonstrate a shift in students' approaches to learning mathematics – from procedural imitation toward deliberate interpretation, planning, and monitoring – directly addressing RQ1 and highlighting the impact of the workshops on learning approaches and understanding. They also indicate *why* this shift occurred, answering RQ2 by identifying the workshop features that students found most beneficial: explicit work on "how to think," opportunities to articulate and critique strategies, structured experimentation that transformed "trial and error" into hypothesis-driven exploration, and tasks that foregrounded concepts before procedures. Theme 1 documents greater confidence and the emergence of reflective problem-solving habits.

6.3.2 Theme 2: Growth in Mathematical Communication and Problem-Solving Confidence

This theme was developed through the combination of initial codes related to collaboration and peer learning, and confidence in mathematical problem-solving. It reflects how the workshop environment helped students move from private, often isolated problem-solving to a more expressive, social, and confident mode of engagement with mathematics. Students' accounts indicate shifts in participation and recognition (Wenger, 1998; Gee, 2000), as well as gains in self-efficacy that are typically achieved through mastery experiences and social persuasion in supportive settings (Bandura, 1997). Through structured group work, peer teaching, and collaborative whiteboard activities, students gained the confidence to explain, critique, and reflect on their mathematical reasoning in ways that extended beyond the technical correctness of answers. This is consistent with research showing that active, dialogic norms foster more profound understanding and persistence (Freeman et al., 2014).

A defining characteristic of this theme was the increased ability to communicate mathematical ideas, both in conversation with peers and in written form. "Zanele" (FG2) noted, "I am able to be confident enough to explain to my group members how to approach problems. And I think that's a really important part that I had to improve in my life." Her reflection illustrates how her confidence and communication evolved together,

particularly in group settings where she had to make her thinking visible and understandable to others. Pedagogically, this reflects movement toward self-regulated learning, as students externalise reasoning, receive peer feedback, and calibrate their strategies (Nicol & Macfarlane-Dick, 2006; Zimmerman, 2002).

For Juan, the shift was more personal and emotional. “Coming from high school, I was never wrong in front of people,” he admitted. “But in the workshops, you are okay with that. You are comfortable with being wrong because you know these people are here to work with you and to help you solve things. They are not here to judge.” His quote reflects a recurring pattern: students felt liberated from the fear of failure and gradually learnt to see error as part of mathematical growth, not a sign of weakness. This re-norming of error as informational rather than threatening aligns with work on productive struggle and belonging conditions under which learners are more willing to risk explanation and revise ideas (Hiebert & Grouws, 2007; Kapur, 2014; Walton & Cohen, 2011).

Another aspect of this theme was the value placed on listening and reciprocal explanation. In one workshop activity, students took turns being the “explainer” and “listener”, which taught them to both articulate and interpret mathematical reasoning. “Hector” (FG4) recalled,

> This other time we had to partner up with someone, and one would speak for three minutes while the other just listened without interruption, then ask questions to show that you were listening intently. That helped me understand how others think and made me better at explaining my own approach too.”

Exercises like this instantiate the disciplinary discourse norms of mathematics reason-giving, precision, and audience awareness while building metacognitive monitoring of one’s own and others’ strategies (Schoenfeld, 2012).

The development of mathematical writing was also central to this theme. Students mentioned that the workshops helped them understand the importance of clean, logical presentation, especially in proofs. Walter reflected:

> In our tests there are bonus marks for good writing, the workshops have helped me make sure whatever I write is presented properly, show all the steps, and where there's a need for explanation, I explain on the side so the person marking can follow.

Together, these accounts suggest constructive alignment at work: teaching-learning activities rehearsed the same justificatory moves and expository standards that assessments later rewarded, clarifying what counts and lowering extraneous uncertainty (Biggs, 1996; Biggs & Tang, 2011).

Importantly, students described this growth not just as an academic skill but as part of their evolving mathematical identity. Confidence in speaking, writing, and explaining became linked to their sense of capability and belonging in a university mathematics environment. Palesa captured this confidence shift in how she now enters public problem-solving spaces:

> [In the workshops] I'm more willing and open to ask questions now…I used to avoid writing on the board unless I was 100% sure, but now I'll start something and the team helps me solve it. There's a silent agreement that nobody is going to be judgemental and everybody's here to learn mathematics and we're being serious about it.

Palesa's narrative reads as a shift from peripheral to more legitimate participation in a community of practice (Wenger, 1998), supported by recognition and shared norms that validate students' voices (Gee, 2000). The workshops appear to have provided the relational and epistemic conditions safety to err, opportunities to explain, clear criteria that

support stable gains in self-efficacy and mathematical communication (Bandura, 1997; Biggs & Tang, 2011).

At the same time, the data reveal tensions that are worth noting. Normalising errors can inadvertently create overreliance on peer validation if not paired with explicit attention to formal standards; similarly, strong alignment with assessment can tip into instrumentalism (students adopting an instrumental orientation to learning, treating tasks as instruments to get marks, not as opportunities to make meaning) if students write for marks rather than for meaning (Biggs & Tang, 2011).

Theme 2 directly addresses RQ1 by demonstrating that students perceived a shift in their approach to mathematics, from private, correctness-oriented work to dialogic practices of explaining, listening, and writing that fostered confidence and a stronger sense of belonging. It also answers RQ2 by identifying the workshop features students considered most beneficial for this change: structured group tasks and reciprocal explanation, explicit attention to mathematical writing and audience, clear criteria that made what counts visible, and a climate that normalised error while inviting reason-giving. These practices did more than polish presentation; they enabled students to see themselves as legitimate participants in mathematical discourse, which in turn supported persistence and willingness to take intellectual risks. Noting the tensions raised, avoiding overreliance on peer affirmation and resisting instrumental “write-for-marks” habits, these findings suggest that communication-rich routines are one of the mechanisms through which the workshops reconfigured students’ confidence and communicative competence.

6.3.3 Theme 3: Reconnecting with Mathematics Through Alignment with Coursework

This theme captures how students experienced the workshops as constructively aligned with lectures, tutorials, and assessments, and how that alignment supported both academic confidence and renewed motivation. In Biggs’ terms, constructive alignment occurs when intended learning outcomes, teaching/learning activities, and assessment tasks are coherently mapped so that what students are asked to do during learning is the

same kind of thinking they will later be asked to demonstrate in assessment (Biggs, 1996; Biggs & Tang, 2011). Students' accounts indicate that, even if not planned as a strict blueprint, the timing, task types, and representational demands in workshops resembled those in tests and exams, creating a sense of preparedness and a deeper approach to learning aligned with course aims (Biggs, 1999; Biggs & Tang, 2011).

Several participants linked this alignment to feeling better prepared, more familiar with expectations, and less anxious about assessments. Aphiwe remarked:

> [The workshops] really helps you figure out how much you know, how much you still need to know, and what you thought you knew but actually don't. It helps you discover more about yourself and how you're actually progressing. Sometimes we even get problems we might not meet in our usual practice; when you work through them you realise, oh, there's this type of problem too, and it's helpful because it prepares you for the test as well.

His observation exemplifies transfer-appropriate processing: practising the same kinds of cognitive operations one will later perform enhances performance on those assessments (Morris et al., 1977). From a cognitive load perspective, alignment reduces extraneous load by making task structures and notational demands predictable, allowing students to allocate working memory to germane processing (schema construction) rather than to deciphering unfamiliar formats (Sweller, 1988; Sweller et al., 2019).

For many, this alignment created a feeling of preparedness and familiarity, which in turn boosted performance and reduced anxiety. Amanda echoed this, saying,

> We got a question that was similar to the one we did in the workshops. It was a breeze going through it; I was smiling through the paper because I was able to do that proof so much that I even got extra marks for good writing. They reinforce good mathematical writing and being able to read mathematics.

Her comment highlights the workshops' role as an academic bridge, reinforcing and clarifying content that students may have struggled with in more traditional, lecture-based settings.

But the impact extended beyond academic support. Several participants described how the workshops helped them reframe their relationship with mathematics, shifting it from something overwhelming to something worth engaging in. “Mbali” (FG6) reflected,

> [The workshops] brought the love back for maths. As much as the content might be difficult, having the desire to do it; that feeling of looking forward to attending the workshops and interacting with different people and getting their perspectives.

Alignment can operate as an affective mechanism: by clarifying what success looks like, it strengthens perceived control and self-efficacy, which in turn reduce anxiety and support persistence (Bandura, 1997; Pekrun, 2006; Biggs & Tang, 2011). Moreover, when tasks feel meaningful and instrumental to assessment, students report a higher utility value and are more likely to reinvest effort, even when the difficulty rises (Hulleman & Harackiewicz, 2009). Participants’ references to looking forward to workshops suggest that belonging and recognition, as well as feeling part of a community engaging in “the real work” of the course, also contributed to renewed intrinsic motivation (Walton & Cohen, 2011; Wenger, 1998).

In direct response to RQ2, students identified the design feature of alignment as one of the beneficial aspects of the workshops for their learning: the timely sequencing of sessions with lectures and assessment-congruent tasks that rehearsed the same forms of reasoning they would later be asked to demonstrate in the assessments.

6.3.4 Theme 4: Facilitator Influence and Supportive Learning Environment

This theme was shaped by codes that captured students’ reflections on the emotional tone, motivational style, and relational qualities of the workshop space, most of which

were attributed directly to the facilitator. Participants consistently described the workshops as supportive, non-judgemental, and emotionally safe. These qualities were not incidental but closely linked to the facilitator's intentional practices, which shaped how students engaged with mathematics both emotionally and cognitively. Students' accounts align with research on psychological safety and belonging as preconditions for risk-taking, articulating reasoning, and refining ideas (Edmondson, 1999; Walton & Cohen, 2011), and with control-value theory[16], which suggests that reduced threat and a stronger sense of control support adaptive engagement (Pekrun, 2006).

The facilitator was a central figure in these accounts. Students described him not only as a mathematics facilitator but also as a mentor and motivator. Karabo (FG6) expressed this clearly:

> It's not even about getting the answers right; it's more about the process of learning how to get to that answer or the approach. There's a meditative session, a bit of reflection, sometimes watching a video in between or taking a break…There's always different things implemented that help us enjoy the learning.

His words highlight how the affective and human elements of the workshop were as impactful as the mathematical content.

"Naledi" (FG4) went further, describing the experience as *"*math therapy". She explained that "The workshops have taught us to be comfortable with being uncomfortable and knowing that it is okay not to get it right the first time; just keep on going; you will eventually get there." This quote reflects a recurring theme in the data: students felt emotionally affirmed through their participation, especially if they had previously experienced anxiety or alienation in mathematics. Normalising error as information, and difficulty as expected,

[16] Control-Value Theory (Pekrun, 2006) holds that students' achievement emotions arise from two appraisals: their perceived control over learning/performance and the value they attach to the task or outcome. Different combinations of control and value predict emotions such as enjoyment, anxiety, or boredom, which in turn influence attention, strategy use, and performance.

supports productive struggle without shame (Hiebert & Grouws, 2007; Kapur, 2014) and can shift appraisals from threat to challenge (Pekrun, 2006).

Many participants highlighted that the workshops promoted vulnerability and courage. Jane (a repeating student), for example, shared that.

> I'm like a person who's really scared of looking stupid and asking questions, so I would rather keep quiet, and I understand that basically led me to failing [last year], so the workshops for me are so helpful because you are not being judged at all when you are in the space. In the workshops, it's like everyone is so involved in helping one another. If you don't understand something, nobody's judging you when you ask a question.

The workshop space became a rare setting in mathematics where not knowing was normalised, and growth was prioritised over performance. This is a characteristic of relational pedagogy (Noddings, 2012) and community-of-practice dynamics in which participation and recognition scaffold more central engagement (Wenger, 1998; Gee, 2000).

Students also recognised that the structure and culture of the workshops encouraged inclusivity. "Andile" (FG3) recalled,

> They did this thing of giving us different whiteboard markers and say they want to see different colours on the board so that they see that everyone's engaging and everyone writes their piece of work there, which is helpful.

These small but intentional design choices helped create a culture of participation, particularly for students who might have otherwise remained on the margins. In equity-oriented pedagogy, such participation structures make contributions visible and distribute opportunities to explain and be heard (Cohen & Lotan, 2014).

The integration of motivational tools also played a role. Students mentioned [the facilitator]'s use of quotes, stories, and even mental health check-ins. "Busisiwe" (FG5) remarked,

> We get so much more mental health help in maths than we do in our own faculty… the fact that they'll take time from our workshops to actually just have a full workshop about how we can improve our mental health, I think that's really, really cool. They care about us as people, not just as students.

These reflective practices elevated the workshops from being merely supplemental instruction to emotionally enriching experiences. Framed through motivation research, such moves can enhance utility value and relevance, promoting students to reinvest effort in challenging tasks (Hulleman & Harackiewicz, 2009), while also signalling care and belonging (Walton & Cohen, 2011).

This theme also intersected with students' broader academic and personal identities. As Naledi commented, "I call it math therapy. It feels like therapy." Her insight underscores the holistic nature of the support students felt, support that extended beyond mathematics and into their sense of academic belonging. In this way, facilitation practices operated both epistemically (clarifying norms of reasoning and communication) and affectively (safety, care), conditions known to foster persistence and self-efficacy (Bandura, 1997; Freeman et al., 2014).

Through the development of this theme, the preliminary codes *Facilitator Influence, Inspirational Guidance, Psychological Support and Emotional Safety* were brought together. Instead of being regarded as distinct categories, they were perceived as two aspects of a relational and emotionally intelligent pedagogy that fosters trust, curiosity, and perseverance.

The data surface tensions that matter for design and scale-up. Strong facilitator centrality can risk over-reliance on a single personality if not paired with routines that cultivate peer-

to-peer agency; similarly, care-forward practices must be balanced with explicit standards so that safety supports, rather than substitutes for, mathematical rigour (Biggs & Tang, 2011). Finally, suppose other course spaces do not share these norms. In that case, students may experience epistemic whiplash, feeling safe and agentic in workshops but reverting to silence in lectures, suggesting the value of course-level coherence in climate and expectations (Wenger, 1998; Biggs & Tang, 2011).

This theme illustrates how the facilitator's ethos and practices transformed the emotional landscape of learning. Students felt guided, inspired, and cared for, which in turn made them more willing to engage, collaborate, and take intellectual risks. In combination with clear norms for discourse and participation, these relational moves appear to have enabled the very behaviours explaining, questioning, and persevering that underpin growth in mathematical understanding. In line with RQ1, these accounts suggest that students' shifts in how they learn, greater willingness to ask, explain, and persist, were made possible by the facilitator's creation of a psychologically safe, inclusive space. In relation to RQ2, participants highlighted the features that helped them the most: a non-judgmental climate and structured participation that made expectations clear and effort worthwhile.

6.3.5 Theme 5: Student Voices on Workshop Hurdles

While student reflections on the workshops were predominantly positive, they also highlighted important critiques, tensions, and unmet needs. Constructed from the initial code Frustrations, Gaps, and Environmental Barriers*,* and the preliminary theme (see Phase 3), Desire for Continuity and Guided Independence*,* this theme captures students' experiences of barriers to full engagement, including structural constraints, emotional and cognitive overload, and challenges in adapting to group-based learning environments. Importantly, students did not merely describe these difficulties; they also offered practical suggestions for improving the workshop experience.

A prominent concern was the physical discomfort of the venue. The learning space, described by many students as hot, crowded, and poorly ventilated, directly affected their concentration and motivation. Karabo explained, “It is very humid and clustered and hot in the venue. So I don’t really like that, and most of the time I just want to get out of the place just to get a breath of fresh air.” “Andiswa” (FG6) echoed this frustration: “Sometimes when you come in there, it’s so hot, and then I start to feel sleepy, and I can’t really focus well.” Such conditions created extraneous cognitive load (Sweller, 1988; Sweller et al., 2019), diverting limited attention resources away from germane processing and undermining the otherwise supportive climate. The tension here is straightforward: the workshops aim to reduce threat and increase engagement, yet environmental stressors reintroduce load and erode the very safety they cultivate.

Time constraints posed another common barrier, particularly when dealing with complex problems that necessitated more discussion and scaffolding. “Lebo”(FG6) shared, “Sometimes we’d struggle with challenging questions, but then time runs out and we don’t finish it. I’m left wondering if my approach was even correct.” His concern was not about “getting the answer” but about closure and validation of the approach. “Khanyi” (FG6) proposed a simple adjustment:

> At the following workshop, before we start with that day’s content, we could look at those questions; may the tutors explain to us how we should approach it, not necessarily give us the answer, but explain to us the way we should approach the question. I think that would be helpful.

Briefly revisiting unfinished problems at the start of the next workshop, with tutors explaining how to approach them without giving complete solutions, keeps the struggle productive (Kapur, 2014) and prevents unresolved confusion from turning into doubt.

Students also highlighted difficulties arising from group dynamics, especially when some group members did not participate equally or were hesitant to speak. Naledi shared, “In the beginning it was really brutal, you get in a group and no one wants to talk, and then

you're the only one talking…and some other person is just yammering on." This points to the well-known tension between cooperative structures and equitable participation: without explicit participation norms and roles, groups can drift into social loafing or unequal status contributions (Cohen & Lotan, 2014). The workshops' broader emphasis on inclusion and listening mitigated this for many, but data suggest that participation structures need to be consistently enacted if collaborative benefits are to be distributed.

Beyond the sessions themselves, students expressed concern about the lack of continuity beyond the first year. Mbali shared, "I would like there to be a MAM2[17] workshop…It's more challenging in MAM2," and worried that without a similar space she'd be going through that on her own and "it could feel overwhelming." Lebo offered a proactive suggestion, "You're on your own and only have one tutorial a week.. so if we all come together again in second year, we can go through the tough questions together."

These comments crystallise a productive tension in guided independence: the workshops successfully scaffold participation and strategy use, but students are anxious about the scaffold fading as they transition to more autonomous settings (Wenger, 1998). Designing for "fading" and "transfer" for instance, by seeding peer networks and norms that outlive the intervention would align with the literature on sustaining belonging and self-regulation (Walton & Cohen, 2011; Zimmerman, 2002)

Another point of tension was the amount and timing of stand-alone wellness sessions. As Karabo put it, "we overdid [the breathing exercises]," especially when "the wellness sessions [are] on such inconvenient times… it feels like I'm just sitting there for two hours when I could have done something better."

Students valued discovery-oriented white board work but wanted to check their method and presentation afterwards. As Andile (FG3) explained, "After working on the board… we also need an ideal way of 'this is what you're aiming for'…this is how you should have approached it."

[17] Second year Mathematics

What distinguishes this theme is the constructive nature of student feedback. Their critiques were not rejections of the workshop model but calls to fine-tune and scale what already works: improve the physical environment to reduce extraneous load; build in brief "return to unresolved problems" segments to stabilise approaches; enact explicit participation roles to support equitable talk; seed peer communities and channels that persist into second year; and avoid same-day pre-test workshops in favour of spaced, retrieval-orientated review. In each case, students' suggestions map naturally onto established principles of cognitive load management, self-regulation, cooperative learning, and assessment-aligned practice.

This theme, therefore, captures the real-world complexities of educational interventions. Even effective, well-received programmes must contend with environmental constraints, evolving learner needs, and the realities of scale and sustainability. The analytic contribution is to show where the workshops' underlying design values safety, alignment, participation, and strategy use can come into tension with venue conditions, time pressure, group variability, and assessment schedules, and how those tensions can be resolved in line with the literature to preserve learning while supporting durable independence.

This theme addresses RQ3 by identifying key barriers to full participation, venue conditions, time pressure/unfinished problems, uneven group dynamics, and the post-year continuity gap as moderators that undermine the very mechanisms the workshops rely on (e.g., safety and participation).

6.4 Summary of Qualitative Findings

This chapter presented five themes developed through reflexive thematic analysis of six focus group interviews with students who participated in the MTWs. The themes directly respond to the study's three qualitative research questions and demonstrate the workshops' multifaceted impact on students' learning. Table 17 summarises how each theme maps onto the questions.

Table 17: Relationship between themes and research questions.

Theme Title	**RQ1**	**RQ2**	**RQ3**
1. Developing Reflective and Strategic Mathematical Thinking	√	√	
2. Growth in Mathematical Communication and Problem-Solving Confidence	√	√	
3. Reconnecting with Mathematics Through Alignment with Coursework		√	
4. Facilitator Influence and Supportive Learning Environment	√	√	
5. Student Voices on Workshop Hurdles			√

Themes 1 and 2 address RQ1 by showing perceived shifts in approaches to learning and understanding. In *Developing Reflective and Strategic Mathematical Thinking*, participants described moving from rote memorisation to deliberate, analytical problem-solving. They learnt to pause, consider problem structures, and select appropriate strategies, indicating growth in metacognitive awareness. In *Growth in Mathematical Communication and Problem-Solving Confidence*, students gained the confidence to articulate mathematical ideas both verbally and in writing. Group work, whiteboard discussions, and opportunities to explain their thinking helped reinforce their mathematical identity and expressiveness. Students' movement from rote memorisation toward more strategic approaches (Theme 1) reflects processes described in APOS theory and aligns with metacognition. Growth in confidence and communication (Theme 2) resonates with social constructivist accounts of learning through dialogue and participation.

Themes 3 and 4 identify what aspects were most beneficial, responding to RQ2. Theme 3 highlighted how the workshops supported conceptual consolidation by aligning closely

with formal coursework and assessment content. This academic reinforcement, in turn, reignited students' interest in mathematics, transforming it from a source of anxiety to one of curiosity and enjoyment. *Facilitator Influence and Supportive Learning Environment* focused on the critical role of the facilitator in shaping an emotionally safe and motivational learning culture. Students credited the facilitator with fostering a space where vulnerability, reflection, and persistence were encouraged, describing the workshops as a place where they felt seen and supported.

Finally, Theme 5 addressed the question of what challenges and obstacles students encountered (RQ3). These included environmental discomfort, limited time for in-depth exploration of complex problems, uneven group participation, and concerns about maintaining independence after the workshop. Despite these barriers, students remained committed to the workshop model and offered practical suggestions for improving its structure.

Together, these themes illustrate the workshops' transformative potential as supplementary academic support and a pedagogical model that integrates cognitive, emotional, and social dimensions of learning. They demonstrate how thoughtfully designed; well-facilitated, and student-centred learning environments can support reflective thinking, build confidence, and foster a sense of community among learners navigating the challenges of university mathematics. While transformative in many aspects, these findings also raise questions about scalability and sustainability, particularly given the reliance on facilitation and the need for alignment with specific coursework. These qualitative insights provide a foundation for interpreting the quantitative patterns observed, and together they inform the integrated discussion that follows in Chapter 7.

CHAPTER 7: DISCUSSION AND CONCLUSIONS

7.1 Purpose and Roadmap

This chapter integrates the discussion and the conclusion of the study. Its purpose is to interpret the quantitative and qualitative findings jointly, articulate the mechanisms that plausibly explain the observed patterns, answer the research questions, and state the study's contributions, implications, limitations, and directions for future work.

The chapter proceeds as follows: Section 7.2 presents an integrated synthesis of the quasi-experimental and thematic results, from which Section 7.3 develops a theoretical account of the mechanisms linking the intervention's design to the observed outcomes. Building on that interpretation, Section 7.4 considers practical implications for course and assessment design, facilitation, and student support, while Section 7.5 gathers methodological reflections and limitations that bound the claims. The answers to each research question are then stated in Section 7.6, before Section 7.7 clarifies the study's theoretical, methodological, and practical contributions. Section 7.8 translates the evidence into recommendations for practitioners, institutions, and students; Section 7.9 identifies priorities for future research; and Section 7.10 closes with final reflections on responsible scaling and the humanised transition into university mathematics.

7.2 Integrated Synthesis of Findings

This section brings together the quantitative and qualitative strands of the study to provide a more holistic account of the Mathematical Thinking Workshops. The integration of these findings follows a convergent, parallel mixed-methods design in which both quantitative and qualitative data were collected and analysed independently before being compared and interpreted in relation to each other. This design enables the integration of complementary insights, where numerical outcomes provide evidence of academic effects, and qualitative narratives add depth, nuance, and context to those outcomes (Adhikari & Timsina, 2024).

7.2.1 Areas of Convergence

A clear point of convergence across strands is that workshop participation is associated with improved conceptual engagement and academic performance in early-semester coursework, alongside shifts in strategy use, metacognition, and confidence.

Students' accounts in *Developing Reflective and Strategic Mathematical Thinking* describe moving from rote procedures to planned, monitored, and multi-strategy problem-solving (e.g., pausing to analyse structure, trying alternative representations, and checking intermediate steps). This corresponds to the quantitative pattern, which shows statistically significant gains on early formative assessments: Test 1F and Test 2F are reliably higher for workshop participants after controlling for multiplicity. By contrast, second-semester class tests (Test 1S, Test 2S) show positive but non-significant estimates, and the estimated effects on the exams and final marks are near zero and not statistically significant. In the higher-order analysis, Test 2F (HO%[18]) shows a positive, unadjusted gain that becomes marginal after Holm adjustment; band-specific estimates indicate a robust HO% gain for the 70–84 NSC band on Test 2F (adjusted pHolm = 0.002), while other HO% effects are directionally positive but not significant. Together, these results align with the qualitative shift toward deliberate, reflective problem-solving, which is closely tied to weekly learning in the first semester.

The theme *Reconnecting with Mathematics Through Alignment with Coursework* captures students' sense that workshop tasks and feedback were well-timed and well-matched to what was expected in assessments. This perceived fit offers a plausible explanation for the strongest quantitative improvements appearing on the first-semester class tests (Test 1F, Test 2F).

Students credit scaffolded practice, group problem-solving, and the supportive facilitation documented in *Facilitator Influence and Supportive Learning Environment* and *Growth in*

[18] HO% = percentage on items requiring non-routine reasoning, structural analysis, and multi-step justification; capped at 100.

Mathematical Communication and Problem-Solving Confidence. These qualitative themes reduced fear of error, willingness to attempt unfamiliar problems, and rising sense of capability, are consistent with the patterned gains on reasoning-oriented work (significant on Test 1F and Test 2F; positive but non-significant on Test 1S, Test 2S, and most HO% estimates after adjustment).

7.2.2 Areas of Divergence

In *Facilitator Influence and Supportive Learning Environment* and *Growth in Mathematical Communication and Problem-Solving Confidence*, students describe feeling more prepared for high-stakes assessments less anxious, more systematic, and more willing to attempt unfamiliar problems. Quantitatively, however, end-of-term examinations and overall course marks show no reliable differences between workshop participants and controls. This divergence suggests that improvements in study practice and confidence may manifest most strongly in proximal[19] coursework contexts, whereas translation to time-pressured, breadth-orientated examinations may require longer consolidation, explicit exam-strategy coaching, or closer alignment between workshop tasks and exam formats.

Students' narratives in *Developing Reflective and Strategic Mathematical Thinking* emphasise analysing structure, trying multiple approaches, and checking intermediate steps, behaviours aligned with higher-order demands. The quantitative strand shows clear, adjusted significance on early formative tests (Test 1F, Test 2F) and directionally positive but non-significant estimates on later class tests (Test 1S, Test 2S). For higher-order (HO%) items, unadjusted positives are observed—with the strongest, adjusted effect in the 70–84 NSC band on Test 2F, yet none of the HO% estimates survive multiplicity correction. Thus, students experience growth in reasoning, but broad, assessment-wide statistical confirmation is limited.

[19] By proximal coursework here I mean assessments that are close to the workshop time, i.e., those that happen during teaching weeks while students are actively attending workshops.

Students often speak as if workshop benefits apply to “everyone who attends”. By design, the quantitative strand draws inferences for specific target populations (e.g., matched ATT for early analyses; overlap-weighted estimates for subgroups with good common support). These estimands intentionally exclude regions of poor overlap. The divergence is interpretive: students generalise from lived experience, while the statistics restrict claims to populations where comparability is credible.

7.2.3 Added Value of Mixed Methods Integration

The integration of qualitative and quantitative findings offered a more comprehensive and multidimensional understanding of the workshops’ effectiveness than either strand could have achieved independently. While the quantitative analyses indicate where and when learning gains are most detectable, the qualitative themes illuminate how and why those gains occurred. The qualitative themes, *Developing Reflective and Strategic Mathematical Thinking*, *Reconnecting with Mathematics Through Alignment with Coursework*, *Growth in Mathematical Communication and Problem-Solving Confidence*, and *Facilitator Influence and Supportive Learning Environment*, suggest a process by which change may occur: a psychologically safe ecology normalises error and inquiry; facilitator practices prompt planning, monitoring, and explanation; and repeated, scaffolded practice crystallises strategies into habits.

This triangulation strengthens the overall credibility of the study and highlights the importance of designing educational interventions that are not only cognitively enriching but also socially and emotionally responsive. The correspondence between students' accounts and measurable gains on early first-semester assessments increases confidence in the interpretation of effects in this context. Moreover, students' voices enrich and contextualise the account, illuminating mechanisms and meanings, rather than validating quantitative results.

Using multiple strands of evidence aligns with best practices in educational research for enhancing credibility and depth of understanding (Creswell & Plano Clark, 2017; Greene

et al., 1989). As argued by Johnson and Onwuegbuzie (2004), combining multiple forms of evidence strengthens the validity of findings and captures the complexity of educational interventions more fully than any single method alone.

As noted in Chapter 4, both the demographic and schooling backgrounds of the workshop cohort provide important context for interpreting these integrated findings. The group included students from a wide range of school quintiles, including those from the most under-resourced public schools as well as those from better-resourced public and private schools. In addition, the cohort reflected South Africa's broader redress categories, which remain closely associated with unequal schooling opportunities. These patterns indicate that the MTWs attracted students who arrived at university with markedly different levels of prior exposure to mathematical resources, pedagogical support, and opportunities to engage with conceptual reasoning. Returning to this earlier discussion, the qualitative and quantitative findings suggest that the MTWs may have been particularly consequential for students whose prior schooling offered limited preparation for the forms of abstract thinking, explanation, and metacognitive regulation expected at university. This reinforces the argument that interventions grounded in conceptual reasoning, collaborative dialogue, and supportive facilitation can play a compensatory role for students most affected by systemic educational inequalities.

7.3 Theoretical Integration and Mechanisms of Impact

The findings of this study offer important theoretical insights into the development of mathematical thinking during the transition from secondary to university mathematics. The results support and extend existing theories of mathematical learning, including Constructivism, CLT, ZPD and APOS Theory, while also contributing to broader discussions about metacognition and inclusive pedagogies in tertiary mathematics education. At the same time, the evidence also reveals limits and tensions with aspects of these frameworks, and this section therefore considers not only points of alignment but also where the data challenge or refine theoretical expectations.

7.3.1 Constructivist Foundations and Active Learning

The workshops were designed with a constructivist orientation, prioritising student activity, sense-making, and peer dialogue over passive knowledge reception. While the qualitative strand did not compare this approach to more traditional, lecturer-centred methods, the qualitative findings were consistent with constructivist intentions: many students described how they shifted from rote memorisation to deeper, conceptual engagement, an outcome consistent with constructivist principles (von Glasersfeld, 1995). This shift was especially evident in the theme *Developing Reflective and Strategic Mathematical Thinking*, where students explained how they began to select appropriate methods based on the nature of a problem rather than applying memorised procedures. Their comments reflected growing awareness of their own learning processes and evidence of constructive abstraction in action.

At the same time, several important critiques temper a simple endorsement of constructivism. First, constructivist designs can underestimate the value of explicit instruction for students who need clear models and guided practice; without sufficient scaffolding, open tasks can increase cognitive load for learners still consolidating foundational understandings. Students voiced this directly, noting that after struggling with hard questions “you’re not given a solution… I want to see if my approach is good or what,” and requesting that tutors “explain how we should approach [the difficult items], not necessarily give us the answer.” Others asked for “a document…explaining each step” to clarify pathways without turning workshops into answer-giving, and some described the open-endedness as disorienting “there’s no boundary… I actually need to be restricted.”

Second, some students may have entered Semester 2 with partially formed concepts or misconceptions, making it harder to connect new material coherently; in such cases, active engagement does not guarantee deep or stable learning. Students described missing conceptual “bridges” between topics, “there’s no bridge between the last topic and this topic… we don’t know the bridges… how to get from this to here”, and

emphasised the need to “build the prerequisite understanding” before tackling more advanced proofs or techniques. At the same time, several students noted that the workshops often *helped* to manage these gaps through facilitator modelling and staged problem progressions, which stabilised ideas when time allowed; however, when debriefs were rushed or postponed, the gaps persisted. Taken together, this points to the need for a deliberately hybrid model that integrates complementary frameworks: constructivist, peer-based problem solving and ZPD-aligned scaffolding; selective, explicit instruction with worked-example fading from Cognitive Load Theory to manage complexity; and metacognitive routines (plan–monitor–evaluate/ICAP) to make conceptual bridges visible while students still do the intellectual work.

While students’ descriptions of increased conceptual engagement and peer-supported problem-solving suggest that the workshops aligned with constructivist principles, the study design did not include a comparison group exposed to more traditional instructional methods. As such, claims about the effectiveness of constructivist approaches must be interpreted with caution.

7.3.2 APOS Theory and Cognitive Development

The findings also map onto the APOS framework (Asiala et al., 1996), which posits that mathematical understanding develops through a progression of mental structures: Action, Process, Object, and Schema. Many students' reflections suggested a movement from the action stage, where rules are applied without reflection, towards the process and object stages, where students could visualise, manipulate, and generalise mathematical concepts. For example, students reported a newfound ability to explain *why* a particular step is used in a proof or to justify a method choice, rather than simply executing it without understanding. They also began to view mathematical entities as objects, which they could analyse, transform, and reason about, hallmarks of cognitive development toward schema-level thinking. These findings suggest that the workshops provided repeated opportunities to advance through APOS stages in a scaffolded, supportive environment.

An illustration arose in collaborative proof work that made interior steps explicit: “we were asked to collaborate on that proof…each individual [wrote] a line… I managed to understand that proof…because everyone was helping me.” This exemplifies interiorisation (coordinating actions into a controllable process) and encapsulation (treating the process as an object for analysing and transformation) (Asiala et al., 1996; Borji et al., 2018), transitions the workshops deliberately aimed to engineer.

Students also described how facilitation and peer explanation supported this shift toward object-like conceptions: “if there were differences… the tutor would… explain why my approach is wrong… and why this is the correct approach, and then… we all come to a common conclusion about why the answer is the way it is.” Others noted that hearing a peer’s framing could unlock a stuck idea: “when someone solved it and I asked why did you do it this way… that was the first time it actually made sense… it just never left my mind.” These accounts align with APOS-consistent growth from action to process to object and the consolidation of schema-level reasoning through repeated articulation, comparison, and justification in collaboration with peers.

Students sometimes wanted more explicit modelling to stabilise developing processes: “maybe… put it on the board… show us how you should have approached it… an ideal way of what we’re aiming for.” Others described moments when open tasks felt under-scaffolded, “there’s no boundary… I actually need to be restricted” or when participation spaces presupposed knowledge (“even the explanation requires you to have knowledge”). These concerns are consistent with Cognitive Load Theory and research on the worked-example effect and faded guidance, which show that novices benefit from explicit models before gradually assuming full responsibility for problem-solving (Chen et al., 2023; Paas & van Merriënboer, 2020; Sweller et al., 2019). Beyond individual cognition, students also emphasised the social mediation of understanding peers “speaking the same language” made reasoning more accessible, pointing to dynamics that APOS, as a fine-grained cognitive lens, does not by itself fully theorise and that are better captured by ZPD-aligned scaffolding and engagement frameworks such as ICAP (Chi & Wylie, 2014).

In closing, the qualitative findings support using APOS as a descriptive core for the cognitive shifts observed, complemented by frameworks that address its blind spots in enactment: ZPD-aligned facilitation for social scaffolding, CLT for sequencing and worked-example fading when prior knowledge is fragile, and metacognitive routines (plan–monitor–evaluate) that make processes visible before they are encapsulated. This blended stance helps explain both the successes (students' growing capacity to justify and reorganise ideas) and the tensions (requests for clearer models and bridges), and it offers a pragmatic route from APOS-inspired analysis to durable, schema-level understanding in first-year university mathematics.

7.3.3 Zone of Proximal Development and Peer Collaboration

The qualitative data strongly supported Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD). The workshops emphasised peer collaboration, exploratory group work, and tutor-guided questioning, all of which are mechanisms that operate within the ZPD, where students can achieve more with guidance than independently (Wood et al., 1976; Rogoff, 1990). Students described precisely this dynamic during proof work: "we were… asked to collaborate on that proof… each individual [wrote] a line… I managed to understand that proof… because everyone was helping me and we were all working together." They also highlighted the value of peer explanations, noting that "sometimes it's better to learn from your peers than from lecturers… we communicate with the same language… it's just easier to get what they're saying and to understand it fully." In practice, this often involved dialogic comparison of methods until a tutor helped the group converge: "first we argue… I want you to convince me why your answer is correct… then the tutor would… explain why my approach is wrong… and… we all come to [a] common conclusion about why the answer is the way that it is." These episodes exemplify movement from other-regulation to self-regulation within the ZPD as strategies are internalised (Mercer & Littleton, 2007; van de Pol et al., 2010). Students reported that working with others allowed them to gain new insights, correct misconceptions, and borrow more efficient problem-solving strategies from peers.

These moments represent learning within the ZPD, where students moved from dependence to the internalisation of skills with the help of a more knowledgeable other.

Importantly, students framed these interactions as social and affective as well as cognitive. They described growing more confident, becoming more comfortable being wrong in front of others, and learning to express themselves mathematically. Such climates of respectful challenge and error-tolerant dialogue are associated with deeper participation and learning in interactive settings (Edmondson, 1999; Brady et al., 2020; Walton et al., 2023). These outcomes suggest that the interpersonal scaffolding provided by both peers and facilitators played a crucial role in shaping their mathematical development. Others noted that rotating roles and prompting quieter peers to take the lead on the whiteboard increased participation and made reasoning visible to the group. This aligns with research on structuring participation so that guidance is widely accessible (Tanner, 2013; van de Pol et al., 2010).

However, it is difficult to define and measure an individual's proximal zone with precision in a fast-moving workshop; miscalibrated tasks can either bore students or overload them (Chaiklin, 2003). Additionally, the quality of scaffolding varies, peers and novice tutors do not always model expert strategies or contingent questioning, so guidance can be uneven unless routines make it explicit (van de Pol et al., 2010). Moreover, ZPD accounts can underplay individual differences in cognitive readiness and prior knowledge; the same prompt can be productive for one student and confusing for another (Kirschner et al., 2006). Finally, ZPD often assumes a socially and emotionally safe environment, which is not guaranteed, particularly for students with prior negative experiences of mathematics or, in the South African context, for students navigating social and economic instability that competes with study and constrains attention and persistence. These conditions can limit access to the very guidance ZPD presupposes (see also student quotes in Section 7.4.2).

7.3.4 Belonging and the Affective Domain

Beyond formal learning theories, the findings point to the centrality of the affective domain, confidence, safety, motivation, and identity in students' mathematical development. Students' narratives indicated that their evolving sense of belonging and willingness to engage were as consequential as procedural or conceptual gains, echoing work showing that emotions and value appraisals shape engagement and achievement in higher education (Pekrun, 2006; Eccles & Wigfield, 2020). The Chapter 6 themes, *Growth in Mathematical Communication and Problem-Solving Confidence* and *Facilitator Influence and Supportive Learning Environment*, illustrate how trust, psychological safety, and relational care created conditions for risk-taking associated with deeper learning (Edmondson, 1999; Brady et al., 2020; Walton et al., 2023).

This emphasis on belonging and identity aligns with identity-as-narrative perspectives in mathematics education, where participation and talk are sites for authoring new mathematical selves (Sfard & Prusak, 2005). In the workshops, students could try on roles such as collaborator, explainer, problem solver, and receive validation from peers and facilitators, a process consistent with research showing that belonging-supportive climates foster persistence and close performance gaps in undergraduate STEM when coupled with authentic opportunities to contribute (Theobald et al., 2020; Laursen & Rasmussen, 2019). South African evidence similarly underscores belonging as a lever for success: low belonging constrains help-seeking and engagement, whereas contexts that communicate care and community promote participation and progress (Sithaldeen et al., 2025; du Toit-Brits, 2022); moreover, sense of belonging predicts intentions to persist among South African university students (Mtshweni, 2024). In transitional settings where prior schooling may have positioned students as underprepared, these identity-affirming experiences appear to be not only ancillary but also foundational to the uptake of higher-order mathematical practices.

7.3.5 Implications for Transitional Mathematics Programs

These theoretical and empirical strands suggest that transitional programmes are most effective when they combine four elements into a single, coherent design: (1) conceptually rich tasks that prioritise sense-making over memorisation; (2) scaffolded social interaction through structured group work and responsive tutor support; (3) metacognitive prompts and reflective dialogue that make thinking visible; and (4) intentional attention to confidence, belonging, and mathematical identity. This profile is in agreement with evidence on inquiry-oriented undergraduate mathematics, where students' engagement in meaningful problems, collaboration for sense-making, and instructor attention to student thinking are central (Laursen & Rasmussen, 2019), and with large-scale syntheses in STEM showing that active, reasoning-centred instruction improves performance and narrows equity gaps (Theobald et al., 2020).

In the context of this study, the workshops succeeded not because they were formulaic or content-heavy, but because they engaged students simultaneously at both cognitive and affective levels. Students described collective proof work in which understanding emerged through shared explanation, an instance of collaborative sense-making that supports interiorisation and transfer. They also emphasised the fit between peer talk and their own language, underscoring the pedagogical value of structured peer instruction in transitional spaces.

Belonging and mutual commitment amplified perseverance. One student captured the ethos with a rowing metaphor, "not for you, for the crew … I don't want to give up on my classmates … and I'm not going to give up on them", highlighting how communal norms sustain effort on complex problems. Others reported identity shifts: "I felt so small … and then I came to the workshops and realised that everyone felt that way … it changed my whole outlook … I'm part of a community programme," a pattern consistent with research linking psychologically safe, belonging-supportive climates to persistence and learning (Edmondson, 1999; Brady et al., 2020; Walton et al., 2023).

Two design choices follow from this evidence and can be enacted without reducing inquiry to answer-giving. First, it helps to structure the interaction rather than the answers: students valued being pushed beyond “answer-getting”, yet many asked for brief clarity on approach at the end of challenging tasks. A short approach focused on consolidations that surface reasoning patterns and comparison with worked exemplars of reasoning (not solutions) can calibrate strategy without undercutting exploration, aligning with research on feedback literacy and internal feedback (Nicol, 2020). This aligns well with Chapter 6 accounts of learning to articulate, critique, and refine methods collaboratively. Second, it is important to design for productive struggle with safety: tasks that invite initial struggle can yield stronger conceptual learning when paired with norms that normalise error and provide timely, least-to-most hints (Kapur, 2014). Students’ descriptions of being comfortable with being wrong and persisting together indicate that the workshops often struck this balance.

In practice, these implications translate into selecting tasks with clear conceptual targets, using equitable participation structures (such as rotating roles and whiteboard protocols) to distribute talk and feedback, and closing sessions with brief, structured debriefs that stabilise ideas while preserving inquiry. These are not first-year-only design principles: students themselves advocated carrying the model forward, “I advocate for collaboration … you really can’t do it alone … let’s get the whiteboards and work together. That’s something I’ll vouch for till this day”, and several asked for a second-year version to buffer isolation in harder modules. Where institutions extend this design longitudinally, transitional programmes can function not merely as on-ramps but as foundations for sustained participation, helping students consolidate identities as collaborators, explainers, and problem solvers.

7.3.6 Contribution to Tertiary Mathematics Education

This study makes a substantive and multi-layered contribution to tertiary mathematics education by showing how a workshop-based intervention can cultivate mathematical thinking, for whom it is most effective, and under what conditions its benefits are

sustained. At its core is an empirically grounded model of support that integrates conceptually rich tasks, structured collaboration, metacognitive routines, and identity-affirming pedagogy. The model does not merely assert that “active learning works”; it specifies design features, planning, monitoring, evaluation cycles, dialogic explanation, facilitation that normalises revision and brief closure routines, that can be enacted in real classes and that map directly onto the assessment constructs where gains are most visible.

Methodologically, the study advances the field by combining rigorous observational causal inference with reflexive qualitative analysis. Using nearest-neighbour propensity score matching with a calibrated caliper, overlap-weighted models focused on the region of common support, cluster-robust inference, and Holm family-wise error control, the quantitative strand provides credible estimates of effect that extend beyond simple correlations. Sensitivity analyses further probe robustness to hidden bias. Set alongside rich focus-group data, these estimates are not left as decontextualised averages: the qualitative strand identifies the mechanisms, metacognitive regulation, collaborative sense-making, and facilitator stance through which effects are likely produced, as well as the boundary conditions, venue quality, pacing, and timing that may explain attenuation in end-of-semester outcomes. In this way, the study demonstrates the distinct value of mixed methods in mathematics education: numbers and narratives are mobilised together to answer, “what works, for whom, and why,” yielding evidence that can be both credible and usable.

A further contribution is measurement innovation. By constructing a higher-order performance index (HO%) from item-level classifications and documenting the distribution of higher-order items across assessments, the study aligns evaluation with its theory of change. This alignment possibly clarifies why impacts concentrate on tests rich in abstraction and justification (for example, Test 2F, which has the highest density of higher-order items) and offers a template for other programmes to audit assessment sensitivity to the kinds of reasoning they aim to foster. In a literature where “higher-order

thinking” is often invoked but seldom measured at the item level, this is a practical step forward.

The study also contributes an equity-attentive lens on transition into university mathematics. Evidence (though limited) that students in the 70–84 NSC Mathematics band exhibited the most apparent benefits on early, reasoning-heavy assessments suggests that well-designed workshops can narrow opportunity gaps without disadvantaging higher-achieving peers. Qualitative accounts of belonging, psychological safety, and collaborative norms, students feel “part of a community programme” and persist “for the crew”, connecting cognitive growth to identity and affect. For institutions where many entrants have historically been positioned as under-prepared, these results suggest that identity-safe, reasoning-centred support can change not only what students know but also how they perceive themselves as mathematical thinkers. Practically, this has consequences for progression: stronger early performance may reduce the likelihood of decanting in extended programmes or attrition, enabling more students to remain on degree-credit pathways.

By articulating a constructivist-plus stance, pairing dialogic, problem-centred activity with selective explicit modelling, cognitive-load-aware scaffolding, and cumulative consolidation, the study contributes to theory refinement. It shows where widely cited frameworks (Constructivism, ZPD, APOS, CLT, ICAP) complement one another in practice and where each has limits unless braided with the others. This theoretically informed pragmatism serves as a necessary counterweight to siloed debates, providing a route from principle to enactment that acknowledges classroom constraints.

At the institutional level, the work translates directly into policy-relevant guidance. It identifies low-cost levers, including venue standards for ventilation, predictable scheduling, facilitator training in contingent questioning and error framing, and session structures that protect time for debriefing, thereby raising the ceiling on impact. It also demonstrates how evidence can be looped back into programme design: in response to student calls for continuity, adapted workshop methodologies have been introduced in

selected second-year modules, providing a concrete model for longitudinal support that sustains gains beyond the first-year transition.

Finally, the study contributes a replicable design and evaluation blueprint. It details tasks, facilitation routines, assessment audits, and analysis pipelines that other departments can adopt or adapt. In a field that often struggles to move from “promising practice” to scalable, evidence-based practice, this project offers both proof of concept and practical guidance. It strengthens the case for holistic, student-centred approaches in STEM education. It shows that, when thoughtfully designed and aligned, workshop-based interventions can deliver measurable early gains, cultivate habits of mathematical thinking, and foster the sense of belonging that underwrites persistence in and through university mathematics.

7.4 Practical Implications

The findings of this study point to several important practical implications for the design, facilitation, and institutional support of Mathematical Thinking Workshops and related interventions in higher education. These implications are especially relevant for departments seeking to support students’ transition into university mathematics, improve retention in STEM pathways, and promote deeper mathematical engagement across diverse student populations.

7.4.1 Curriculum Alignment and Student Support Services

Students appreciated it when workshop topics aligned with their live coursework, especially in the weeks immediately preceding class tests. Several described a sense of recognisable structure “I was smiling through the paper” when the representational moves and problem types in workshops mirrored those in assessments. This enthusiasm speaks to the power of curriculum alignment in making learning feel relevant, timely, and achievable for students who are still building confidence. Alignment also lowers extraneous cognitive load: when tasks echo the conceptual territory students are

traversing in lectures, they can spend more time reasoning and less time guessing what matters.

That said, alignment is not an unqualified good. A tightly coupled schedule can slide into instrumental alignment, a form of constructive alignment that steers workshops toward rehearsing likely test items rather than cultivating durable habits of explanation and strategy. When this happens, students can view workshops as a shadow lecture or test-prep session, which narrows their ambition, encourages pattern-spotting over sense-making, and reduces opportunities for the kind of struggle that strengthens transfer. The quantitative results plausibly underscore this tension: effects were most substantial on early assessments with a high density of higher-order items, yet they attenuated on summative exams that sampled a broader construct. Over-focusing workshops on immediate assessment coverage may therefore improve near performance while starving the very capacities needed later. There is also a practical risk: lecture timetables shift, topics overrun, and unanticipated cancellations occur. If the workshop pipeline is yoked too rigidly to lecture pacing, misalignments increase, leaving some sessions underprepared and others rushed.

A more robust stance is principled alignment rather than tight topic matching. In practice, this means anchoring workshops to a published “construct map” that names the forms of reasoning students must demonstrate (e.g., generalising, definition use, argument structure, representation switching), then selecting tasks that are adjacent to rather than copies of recent lecture material. When sessions open with a two-minute “why this now?” that links the task to course goals and close with an “approach summary” that makes the structure explicit, students still experience relevance without workshops collapsing into test rehearsal. This approach also supports equity: students who miss a lecture due to transport or family responsibilities can still participate meaningfully because the workshop foregrounds reasoning patterns, not a narrow slice of content from a particular day.

Constructive alignment, in its best sense, asks that assessments sample the very thinking workshops cultivate. The item-level audit revealed that not all assessments did so equally.

Where higher-order items were dense, workshop effects were visible; where assessments emphasised memorisation or procedural breadth, effects were muted. The implication is two-sided. In workshop design, workshops should periodically simulate high-stakes conditions so that students practise recruiting explanation and checking under time pressure. In course assessment design, assessments should include a transparent proportion of higher-order marks and explicit rubrics for explanation and justification, so that students recognise the value of the workshop habits and are rewarded for them. Without this reciprocation, constructive alignment risks becoming a one-way street in which support spaces do the intellectual heavy lifting but summative assessments do not credit it.

Alignment also has an affective dimension. Students voiced that they valued workshops as spaces where they could ask “basic” questions, be wrong in public, and learn the social norms of mathematical talk. When alignment choices crowd out these relational functions, by overpacking sessions with “coverage,” or by bolting on wellness content at moments when students are under acute deadline pressure the social contract frays. As one student put it, “the psychology part… we overdid it… and we’re under too much pressure from academics”; another noted that wellness sessions were “at inconvenient times… I could have finished my lab report by now, but I have to sit here for two hours”; a third balanced both views: “mental health is important… there’s a time when the wellness workshops are amazing… but in tight weeks we can’t afford to lose two hours I still have a quiz to study for.” The remedy is not to abandon belonging and wellbeing, but to integrate them lightly and context-sensitively: micro-interventions that take minutes (e.g., normalising statements at the outset, quick check-ins), coupled with predictable referral pathways for students who need more support. In short, let workshops remain primarily academic spaces that are also identity-safe, rather than trying to incorporate full wellness content during peak assessment weeks.

Finally, alignment should be longitudinal, not just week-to-week. The students requested that the model continue into the second year; this work has now begun. Extending the exact principled alignment reasoning-dense tasks, equitable participation norms, short

approach-focused debriefs into later modules creates continuity that stabilises gains and makes transferability across contexts more likely when abstraction deepens. Departments can formalise this by convening a small curriculum-workshop working group that meets each term to review assessment constructs, examine item-level performance, and adjust the construct map. In that forum, alignment ceases to be a scheduling exercise. It becomes an ongoing conversation about what kinds of mathematical thinking the programme values and how teaching, support, and assessment will jointly make them visible and valuable to students.

7.4.2 Scalability and Sustainability

If this model is to transition from a promising pilot to a durable feature of the mathematics curriculum, three interlocking strands must be designed in tandem: people, structures, and data. The people strand concerns facilitator capacity and culture. The workshops live or die on the quality of facilitation; content expertise alone is not sufficient. A sustainable model develops a pipeline of tutors, graduate students, and early-career lecturers who are selected for a student-centred stance and then inducted into a familiar playbook: contingent questioning (least-to-most hints), error framing that normalises revision, equitable talk routines, and basic cognitive-load management. Co-facilitation (pairing a novice with an experienced facilitator) and brief rehearsal “micro-teaches” before each teaching block help maintain quality as the programme grows. Light-touch observation and feedback cycles ten minutes of notes on who spoke, who wrote, and what forms of reasoning were surfaced, create a coaching culture without adding heavy evaluation burdens. This addresses the most acute scalability risk: staff turnover or rapid expansion, which can dilute the very pedagogy that makes the workshops effective.

The structures strand ensures the programme remains visible and valued as it grows. Workshops are sustainable when they are timetabled, credited, and resourced not when they are framed as optional extras. Departments should ring-fence predictable times and venues with good ventilation. Integration into course structures, credit-bearing participation, and assessment rubrics that reward explanation and justification.

Messaging to students is crucial: onboarding should explicitly position workshops as spaces for collaborative growth and strategic problem-solving, not as remedial clinics. Students highlighted how current advising can inadvertently stigmatise participation: "there is a stigma attached to it because the student advisors are saying if you had above 85 [for high school Mathematics], you don't need to take the workshops." This perception is understandable given the recruitment practice in which advisors helped identify at-risk students, and those with an NSC Mathematics score of ≥85% were typically not invited to the workshops, a threshold that can tacitly signal that workshops are only for the underprepared. To counter this, departments should adopt a consistent narrative from lecturers and facilitators, "this is where we practise how to think mathematically, aloud and together", and consider structures that broaden access (for example, cohort-wide credit or explicit pathways for high-achieving students), so that early misconceptions about purpose do not depress engagement as cohorts scale.

The data strand closes the improvement loop. Programs that scale sustainably make their evidence needs modest and routine, tracking attendance alongside simple participation proxies (e.g., completion of role rotations; number of groups reaching a brief closing synthesis). This "dose–response" view highlights topics that need stronger scaffolds, rooms where participation is uneven, and sessions where closure slipped. Small dashboards shared with course conveners keep alignment honest and enable quick adjustments before problems calcify at scale. As capacity grows, a shared repository of tasks, annotated reasoning exemplars, and short approach summaries from those closing syntheses prevents constant reinvention and preserves fidelity across multiple sections.

A candid look at what did not scale well in the workshops setting points to concrete fixes. Where sessions ended without closure, the absence of solutions, deliberate, to de-centre answer-getting, left some students uncertain about approach. The scalable remedy is architectural, not philosophical: build in an 8–10-minute whole-group debrief that surfaces the structure of one or two hard problems and post a short "approach summary" afterward. Where activities felt under-scaffolded for students still consolidating prerequisites,

introduce brief model-based mini-lessons at the start of a block (worked-example reasoning with faded guidance), and keep a bank of just-in-time micro-lessons for prerequisite repairs. Where talk was uneven, make equitable participation a non-negotiable routine (rotating roles, “everyone writes” interludes, rotating public scribes); these are easy to scale and inexpensive to maintain. Where venue heat, noise, or overcrowding eroded concentration, set minimum venue standards as a matter of policy rather than goodwill; no workshop pedagogy can compensate for chronic environmental load. Finally, where effects attenuated on summative exams, insert brief, time-bounded “exam-style” interludes in later weeks and make explicit the transfer from workshop strategies to the constructs examiners actually score.

7.5 Methodological Reflections and Study Limitations

This section reflects on the methodological decisions made throughout the study and highlights the strengths and limitations of the research design. It also considers the implications of the researcher’s positionality and reflexive engagement with the data.

7.5.1 Strengths of the Mixed Methods Design

A key strength of this study is its convergent mixed-methods design, which brought quantitative estimates and qualitative explanations into the same analytic space. Rather than treating numbers and narratives as parallel tracks, the study integrated them to address the classic questions of educational evaluation, what worked, for whom, and why, yielding meta-inferences that neither strand could have produced alone. This integration was planned at the design stage and enacted at the analysis stage to enhance interpretive power and practical usefulness.

On the quantitative side, the design strengthened causal interpretability in an observational setting. Rich pre-treatment covariates (e.g., NSC/NBT indicators) supported nearest-neighbour propensity-score matching with a calibrated caliper to secure balance between attendees and non-attendees. At the same time, overlap-

weighted models targeted effects in the region of common support. Cluster-robust standard errors accommodated control reuse, Holm family-wise error control tempered false positives across outcome families, and Rosenbaum-style sensitivity checks probed the robustness of key findings to modest hidden bias. Together, these choices move the analysis beyond simple associations toward more credible estimates of the effect of attending the workshops on outcomes that matter for students and departments.

Measurement was aligned to the theory of change. The study constructed a higher-order performance index (HO%) based on item-level classification of problems that require transfer, justification, and multi-step reasoning, with transparent rules for capping and inclusion. By documenting the density of higher-order items across assessments (e.g., the notably HO-rich Test 2F), the evaluation linked where effects emerged to what, precisely, was being measured. This item-level audit is an advance on course “overall mark” analyses that can mask changes in the kind of thinking programmes aim to cultivate.

On the qualitative side, the study added necessary depth and mechanism. Focus-group interviews timed across the academic year captured students’ evolving perceptions of the workshops’ cognitive, social, and affective affordances. A reflexive thematic analysis surfaced stable themes (e.g., alignment with coursework, collaborative sense-making, facilitator stance, belonging and agency) as well as counter-themes (e.g., calls for clearer modelling and closure late in the year). The focus group format enabled participants to build on one another’s accounts, generating dialogic explanations (how students planned, monitored, and evaluated their work; how they distributed roles; how norms around error and risk-taking were negotiated) that would likely have remained tacit in one-to-one interviews. Credibility was enhanced through an audit trail of analytic decisions, thick quotation to ground claims, and attention to negative cases where experiences diverged from the central pattern.

The integration of strands was itself a strength. Quantitative gains on early assessments and higher-order items were explained by qualitative evidence of metacognitive routines,

identity-safe collaboration, and facilitator modelling; conversely, the attenuation of effects on summative outcomes was illuminated by students' accounts of timing and rising assessment pressure. Where strands diverged, the design allowed for the interrogation of why (e.g., strong qualitative endorsement of benefits versus weaker signals on finals), producing bounded conclusions that specify the conditions under which impact is most likely. This mixed-methods logic triangulation for credibility, complementarity for mechanism, and development for optimisation supports actionable recommendations for design, facilitation, and policy.

Finally, the study's positionality and reflexivity were treated as methodological resources rather than threats to validity. The researcher's insider knowledge of facilitation practices helped formulate plausible mechanisms and probing questions (the questions that the external interviewer asked were designed by the researcher (myself), while explicit prompts for critique during focus groups and the inclusion of disconfirming evidence tempered advocacy. This stance increased transparency about how interpretations were made and why some design decisions (e.g., withholding complete solution keys to de-centre answer-getting) had both pedagogical value and evaluative consequences that required architectural responses (e.g., protected debrief time).

In summary, the mixed-methods design provided credible estimates, mechanistic insights, and practical guidance. It connected measured changes in performance to observed changes in how students approached mathematics, specifying boundary conditions for effectiveness. It generated a replicable evaluation blueprint, strengths that collectively justify the study's claims and its usefulness for departments seeking to design, scale, and sustain workshop-based support in university mathematics.

7.5.2 Limitations and Constraints

Despite its strengths, the study has several limitations that qualify the claims that can be made. The design is observational: propensity score matching and related methods can balance observed pre-treatment covariates and reduce bias, but they cannot rule out

unmeasured confounding. Factors such as motivation, help-seeking orientation, time-on-task, prior experiences with group work, or differential access to informal peer support may have influenced both workshop attendance and outcomes. The use of balance diagnostics, overlap weighting, cluster-robust inference, Holm family-wise error control, and sensitivity checks enhances credibility; however, causal statements should still be interpreted cautiously.

There are also constraints on the estimands and on generalisability. Effects are reported for attendees (ATT) from the matched design and effects for the overlap population (ATO) from overlap-weighted models. ATT describes the benefit to those who attended; ATO describes effects for students with moderate propensities to attend (i.e., those on common support). Neither quantity automatically generalises to students at the extremes of propensity. Moreover, nearest-neighbour matching with a caliper discards some cases (trading external validity for internal comparability), while overlap weighting down-weights extremes (shifting the target population). These effects should not be treated as “the effect for all students”.

Model dependence is another limitation. Propensity scores are estimated, not known; alternative specifications, link functions, covariate sets, interactions, could alter who is deemed comparable, thereby affecting both the matched sample and the weights. The caliper choice (0.2 SD of the logit) is defensible but not unique; tighter calipers can improve balance at the cost of sample size and power. Although post-matching balance was good, residual imbalance on unmodelled interactions may remain. Cluster-robust standard errors help with reused controls but can be variable-hungry; with modest samples, estimates may still be noisy. Holm adjustment limits false positives across families but reduces power, especially for interaction tests and HO% comparisons, raising the risk of false negatives.

Measurement choices also impose limits. The higher-order outcome (HO%) was derived from an item-level audit intended to capture transfer, multi-step reasoning, and justification. Despite documented inclusion/exclusion rules and capping, item

classifications were not independently double-coded, so inter-rater reliability was not established. The number and difficulty of HO items varied across assessments (for example, Test 2F was HO-dense), meaning HO% is not equated across instruments. The cap at 100 introduces the possibility of ceiling effects, and without an equating framework, cross-test comparisons assume construct comparability that was not formally verified. More generally, overall marks amalgamate heterogeneous skills; null effects on finals may partly reflect construct misalignment rather than a lack of learning.

Uncertainties surrounding exposure and implementation further complicate the interpretation. Dose–response was not modelled (e.g., number of workshops attended or within-session participation), so intensity and quality of engagement cannot be linked to outcomes. Implementation likely varied with facilitator style, room conditions, group composition, and week-to-week timing; such heterogeneity introduces noise and threatens internal validity. Potential spillovers, workshop students helping non-attendees, would violate the no-interference assumption and bias estimates toward zero. Concurrent supports (tutoring, lecturer consultations) were not systematically tracked, so co-intervention effects cannot be excluded.

Contextual factors may also confound impacts. Students repeatedly cited venue heat/noise and end-of-semester workload as barriers. Since these conditions vary non-randomly across sessions, they could depress outcomes differentially and complicate attribution. The strong early-semester signal, coupled with weaker exam effects, may therefore blend pedagogical impact with ecological constraints.

The qualitative strand carries its own limitations. Focus groups recruit volunteers; self-selection may bias accounts toward more engaged or satisfied students. Group dynamics can mute dissent (conformity, deference to confident peers) and induce social desirability, particularly when discussing facilitation. Although critique was actively invited, non-attenders and fully disengaged students are under-represented. The sample was not banded by NSC level, so narratives cannot be attributed to achievement subgroups. Quotations document perception and meaning making, not frequency or effect size.

My positionality is both a resource and a risk. As a designer-facilitator and analyst, I had insight into mechanisms but also a stake in the programme's success. Reflexive journaling, explicit prompts for critique, triangulation with quantitative patterns, and the inclusion of counterevidence were used to temper confirmation bias; yet, interpretation inevitably reflects a situated lens.

Scope and context further bound external validity. This is a single-institution, single-cohort evaluation within a specific national and departmental setting. Resource constraints (facilitator availability, room quality) and institutional calendars shape what is possible; replication elsewhere will likely require adaptation. The analysis emphasises short-term academic outcomes; cohorts were not followed longitudinally to test durability into second-year courses.

Finally, several analytic choices were pragmatic. The analysis plan was not pre-registered; model selection and outcome-family definitions were theory-driven but remain subject to researcher discretion. Interaction tests (e.g., Treatment × Band) are underpowered in modest samples, so wide intervals mean meaningful moderation may have been missed. Conversely, multiple testing increases the chance of Type S/M errors, wrong-sign or exaggerated-magnitude estimates, for more minor effects.

7.5.3 Reflection on Data Collection and Analysis

Six focus group interviews proved fruitful in generating detailed, context-rich data. Participants engaged with prompts, and with one another, in ways that revealed not only what they thought but also how they reasoned about their experiences. The decision to use reflexive thematic analysis (RTA) was well-suited to this aim. RTA enabled flexible yet rigorous theme development with an explicit emphasis on meaning-making, interpretation, and the active role of the researcher in constructing themes (Braun & Clarke, 2006, 2019, 2021).

Data collection and analysis proceeded iteratively. Early patterns were tested and refined through multiple cycles of coding and memoing, and themes were shaped through sustained engagement with the full dataset, rather than relying on isolated “headline” quotations. Throughout the process, I examined internal homogeneity (coherence within a theme) and external heterogeneity (clear boundaries between themes), continually checking alignment with the study’s research questions. An audit trail of analytic memos, decision logs, and saved code iterations to support transparency.

The same characteristics that made RTA powerful in this context also introduce limitations that qualify the claims that can be made. Because RTA centres the researcher’s interpretive agency, subjectivity is a constitutive feature rather than a nuisance parameter. This fosters analytic insight but raises the possibility of confirmation bias or selective attention if reflexivity is weak. I mitigated this risk through structured, reflexive journaling, actively seeking negative cases that challenged emerging interpretations, and discussing tentative themes with three supervisors (two of whom were not involved in the MTWs design and implementation), while also checking for bias or selective attention. Nevertheless, another analyst working reflexively from a different theoretical lens might have produced alternative but legitimate accounts. RTA also eschews inter-rater reliability indices as quality markers on philosophical grounds (Braun & Clarke, 2019, 2021). Readers accustomed to codebook or positivist traditions may therefore experience less comfort with auditability; in this paradigm, coherence, transparency, and reflexive depth replace kappa[20] statistics as indicators of quality.

RTA’s focus on the quality of meaning rather than frequency means that the analysis is not designed to estimate the prevalence of a view across the population. I do not, for example, claim that a certain percentage of students desired more approach-focused closure; if stakeholders require prevalence estimates, a complementary survey would be needed. The method’s flexibility can also tempt analysts into topic summarising that

[20] Kappa statistics (κ) measure inter-rater reliability, the extent to which two or more coders agree on categorical judgements in qualitative analysis, while adjusting for chance agreement; in this study, which used reflexive thematic analysis (RTA), κ was not computed because RTA treats coding as interpretive rather than consensus-seeking (Cohen, 1960; Viera & Garrett, 2005; Braun & Clarke, 2019).

mirrors the interview guide rather than developing interpretive accounts. I sought to avoid this pitfall by anchoring themes in the study's theory of change, metacognition, collaborative sense-making, and belonging and by writing analytic narratives that link data fragments to explanatory claims.

These methodological considerations intersect with the focus-group format. Group dynamics can mute dissent through deference to confident peers or social desirability pressures, and volunteers may over-represent more engaged or satisfied students. While inclusive facilitation methods were employed (such as turn-taking prompts and invitations to counterexamples), these dynamics remain a potential source of bias that RTA's strengths do not eliminate. In terms of the scope of conclusions, the analysis provides analytical generalisation and theoretical insights that readers can apply to similar contexts, rather than statistical generalisation. Finally, high-quality RTA is time-and skill-intensive, requiring immersion, iterative rewriting, and careful evidencing with quotes, which can be challenging within institutional timelines. This meant that extensions, such as member consultations, were outside the scope of this cycle.

7.5.4 Reflexivity and Researcher Positionality

As one of the designers and a supporting facilitator of the Mathematical Thinking Workshops, I occupied a dual role that carried clear advantages and equally clear risks. Insider status provided a granular understanding of the intervention's pedagogical intentions, task design, and institutional constraints, and it helped establish rapport during the recruitment process. I recruited students to participate in the focus groups, which likely encouraged familiarity and trust. To mitigate the effects of power and desirability, however, I did not conduct the interviews; a trained, independent interviewer facilitated all focus group sessions. Even so, proximity to the programme heightened the possibility of seeing what I hoped to see confirmation bias, selective attention to supportive evidence, or an overly charitable reading of ambiguous cases.

In keeping with reflexive thematic analysis, I treated subjectivity not as a contaminant to be eliminated but as an analytic resource to be worked with deliberately. Throughout the project, I maintained a reflexive journal to surface assumptions before and after interviews, and recorded decision trails for coding and theme development. I actively sought negative cases, instances that complicated or countered emerging interpretations. Although I did not conduct the interviews myself, the independent interviewer opened each session by emphasising voluntariness, confidentiality, and the value of critical feedback, using prompts that explicitly invited dissent and difficulty (e.g., “What wasn’t helpful here?” and “When did this approach break down?”). I designed the interview guide and briefed the interviewer to foreground these norms; these practices did not remove subjectivity, but they situated it and made it answerable.

Positionality also shaped what could be known. My involvement in workshop design likely influenced the questions I prioritised and the constructs I foregrounded (e.g., metacognition, collaborative sense-making, belonging). It may have introduced a halo effect in some groups, students expressing appreciation they thought I wanted to hear, or a chilling effect for others who hesitated to criticise practices associated with a member of the research team. To counterbalance this, I anonymised all quotations at the point of transcription, included countervailing excerpts in the write-up, and read qualitative claims alongside the quantitative strand rather than using one to vindicate the other. Where patterns aligned (e.g., early gains on higher-order items and students’ descriptions of planning, monitoring, and explanation), I reported convergence; where they diverged (e.g., strong perceived value but weak effects on finals), I reported tension and examined contextual explanations such as timing, venue conditions, and workload.

Since I am embedded in the department, I am also entangled in the power relations of the setting. Participants may still have perceived me as part of the instructional hierarchy. I attempted to mitigate this by inviting an independent interviewer who adopted a facilitative stance in interviews, listening more than explaining, validating uncertainty, and redirecting praise toward specific practices rather than individuals. However, I cannot guarantee that all participants felt equally free to speak. This recognition informs the

boundedness of my claims: the account presented here is a reasoned, transparent interpretation from a situated vantage point, not a neutral view from nowhere.

Finally, reflexivity influenced how findings were translated into implications. For example, my advocacy for approach-focused debriefs and “worked reasoning” exemplars reflects both what students have asked for and my own pedagogical commitments; I make those commitments explicit so readers can evaluate the recommendations with a clear understanding. In short, the study’s methodological stance acknowledges that knowledge here is co-constructed at the intersection of students’ experiences, institutional realities, and my interpretive labour. By integrating multiple data sources, documenting analytic choices, and remaining candid about proximity and bias, I aim to offer a trustworthy account that is both methodologically sound and humanly attentive to how learning unfolds in this context.

7.6 Answers to the Research Questions

This section addresses the study’s six research questions by interpreting the findings from both the quantitative and qualitative components of the research. The integrated analysis reveals how the MTWs influenced students’ cognitive strategies, affective experiences, and academic outcomes. Each subsection interprets findings in direct relation to a research question.

7.6.1 Quantitative Research Question 1

To What Extent Do Mathematical Thinking Workshops Impact Students’ Academic Performance When Controlling for Pre-University Academic Achievement?

This research question was addressed using the Propensity Score Matching (PSM) approach, which aimed to isolate the effect of attending the Mathematical Thinking Workshops on first-year university mathematics outcomes by controlling for pre-existing differences in academic preparedness. Specifically, students were matched on key pre-university covariates to estimate the Average Treatment Effect on the Treated (ATT). This

strategy provided a more reliable comparison between workshop participants (the treatment group) and non-participants (the control group).

The findings indicate that the workshops had a statistically significant, positive effect on students' performance in early-semester tests, specifically Test 1F and Test 2F, as adjusted for Holm. These outcomes align with the workshop's focus on foundational mathematical thinking, conceptual understanding, and strategic problem-solving skills, elements which are often more prominent in early formative assessments. The ATT for Test 2F, for instance, was notably higher, possibly indicating that the benefits were evident across multiple early assessments within the first semester. This pattern is consistent with the composition of higher-order items across assessments: Test 2F contained the highest number and density of HO items (6 of 7), compared with Test 1F (3 of 7), Exam F and Test 1S (3 of 6 each), Test 2S (2 of 6), and Exam S (3 of 10). The concentration of HO items in Test 2F likely increased its sensitivity to the kinds of strategic and conceptual reasoning fostered by the workshops, amplifying detectable gains on that assessment.

However, the analysis also indicates an absence of significant treatment effects of second-semester class tests, final examinations, and overall semester marks. The positive impact of the workshops may diminish over time if students do not continue to engage with the strategies introduced, particularly as the content becomes more advanced and cumulative. Differential student engagement could also play a role, with participants perhaps being more motivated during the initial part of the workshops but experiencing a decline in engagement as the semester progressed and academic pressures increased. As one student put it, "But all the psychology, I feel like we overdid it. And we are having too much pressure from our academics. I would like the workshop to be flexible…to align with our needs." foregrounding how workload and timing can sap motivation and reduce uptake of workshop strategies later in the term. Another participant highlighted a scheduling conflict:

> Sometimes the wellness sessions [are] at such inconvenient times… I had a lab report due that Friday. It just feels like I could have finished my lab report by now, but now I have to sit here for two hours,

underscoring how rising assessment pressures can crowd out engagement. These comments are drawn from Focus Group 6, conducted at the end of the second semester's last term (17 October 2023), when workload pressures are typically at their peak.

The absence of significant effects on final exam performances and cumulative scores (final marks) mirrors patterns reported in the literature, where the impact of interventions often tapers as courses progress and content becomes more complex (Wan et al., 2021). This “fade-out” pattern is well-documented in education broadly and has been observed within university courses where early-term gains fail to persist to summative exams (Bailey et al., 2020; Freeman et al., 2007).

The quantitative findings are complemented by qualitative data, in which students consistently reported increased clarity, strategic thinking, and confidence in coursework. Participants described how the workshops helped them approach problems more reflectively and collaboratively. For example, some noted that concepts explored in the workshops directly mirrored those found in assessments. “I was smiling through the paper,” one student recalled, highlighting a perceived link between workshop content and academic success.

The results suggest that the Mathematical Thinking Workshops had the most impact during the early stages of the semester. While the lack of significant effects on later assessments may indicate the need for continued support throughout the semester, the positive early gains suggest that the workshops’ value lies in their effectiveness as an intervention for helping students navigate the critical transition into university mathematics. There is a strong possibility that previously, many of the students in the cohort would simply not have made it through to the second semester. The fact that they performed so much better in the first-semester tests indicates that they may have been

decanted (moved into an extended programme) and thus had to transition into a completely different stream. Instead, they continued, and while their results were not much higher than similar students at the end of the year, likely, many of the workshop students would not have made it through to the end of the year.

7.6.2 Quantitative Research Question 2

How Does the Impact of the Workshops Vary Across Different Levels of Pre-University Mathematics Achievement?

To examine heterogeneity by prior preparation, the analysis stratified students into two NSC Mathematics bands (70–84 and 85-100). Band-specific effects were estimated using overlap-weighted models targeting the Average Treatment effect for the Overlap population (ATO) with robust standard errors. Multiplicity was controlled via the Holm adjustment across the two band tests per outcome. A pooled Treatment × Band interaction was assessed to determine whether effects differed reliably between bands.
For the core course outcomes, the first-semester class tests showed large, statistically significant workshop effects in both NSC bands (Test 1F and Test 2F). In contrast, second-semester class tests (Test 1S and Test 2S) yielded positive but non-statistically significant results in both bands, and the final examinations and overall marks were near zero and not statistically significant. Treatment × Band interactions for these outcomes were non-significant, indicating no reliable moderation by NSC band on the assessments where the workshops demonstrably helped (the early class tests). This broad benefit without strong moderation is consistent with large-sample syntheses showing that active, reasoning-orientated instruction raises performance across the achievement spectrum in undergraduate STEM (Freeman et al., 2014).

For higher-order items (HO%), the clearest signal is concentrated in the 70–84 band on Test 2F, with a large and statistically significant gain (+19.03 percentage points; pHolm = 0.002). Other HO% estimates in 70–84 is directionally positive but not significant after Holm adjustment, and HO% effects in the 85-100 band are generally more minor and not significant. The Treatment × Band interaction for HO% (e.g., on Test 2F) does not reach

conventional significance, so the contrast is suggestive rather than definitive. Importantly, Test 2F also had the highest density of higher-order items in the suite of assessments (6 out of 7 long-answer items), which makes it especially sensitive to gains in strategic problem solving and proof-based reasoning cultivated in the workshops. The item audit underlying this classification is documented in Appendix G, which lists all included/excluded Section C questions used to compute HO%.

The pattern of stronger HO gains for the moderately prepared 70-84 group on the reasoning-rich assessment aligns with evidence that student-centred, active approaches disproportionately support students with more room to grow, narrowing performance gaps without harming higher-achieving peers (Haak et al., 2011; Laursen et al., 2014; Theobald et al., 2020). Active-learning designs that emphasise frequent practice with higher-order cognitive skills show the most considerable benefits for students who start lower on prior measures, a mechanism plausibly at play here.

Because NSC Mathematics did not band the focus-group sample, the qualitative strand cannot directly answer this moderation question; participants' narratives cannot be attributed to the 70–84 or 85-100 subgroups. Nonetheless, the qualitative themes (strategy uptake, metacognitive monitoring, alignment with coursework, and growing confidence) provide a plausible account of why both bands benefit on early assessments and why the most visible higher-order gain appears in the 70–84 band on Test 2F. This interpretation is also consonant with work showing that explicitly teaching metacognitive regulation in university STEM can amplify transfer to non-routine problems (Stanton et al., 2021), the very behaviours indexed by the HO% measure. Overall, the workshops operate as broad-based support for early class tests across achievement levels, with targeted evidence of higher-order gains for students in the 70–84 band and limited statistical moderation elsewhere.

7.6.3 Quantitative Research Question 3

To what extent do the Mathematical Thinking Workshops impact students' performance on higher-order assessment items in first-year university mathematics?

Higher-order performance was operationalised as the percentage of points earned on items requiring non-routine reasoning, structural analysis, and multi-step argumentation (HO%), with scores capped at 100 where necessary. Consistent with Chapter 5, HO effects were estimated as ATTs using the matched sample from the propensity-score matching design (1:1 with caliper; cluster-robust standard errors), and family-wise error was controlled with Holm adjustment across the six assessments. In interpreting these effects, it is important to note the assessment composition: Test 2F contained the highest density of higher-order items (6 of 7), exceeding Test 1F (3 of 7), Exam F and Test 1S (3 of 6 each), Test 2S (2 of 6), and Exam S (3 of 10). Given constructive alignment principles (Biggs, 2014), an assessment richer in higher-order targets is more sensitive to gains in the specific forms of metacognitive skills the workshops cultivate.

Across assessments, point estimates are generally positive for workshop participants on higher-order items, with the clearest signal on Test 2F; however, assessment-level HO effects do not remain statistically significant once multiplicity is controlled. Later class tests and the examinations yield small, imprecise, and non-significant HO differences. One plausible explanation for this is that the workshops emphasised interactive and constructive engagement (peer explanation, modelling of thinking aloud, and reflective prompts), the forms of participation most strongly linked to deeper learning and transfer in the ICAP[21] framework; thus, effects should be most visible where assessment tasks elicit those behaviours (Chi et al., 2018; Chi & Wylie, 2014).

The qualitative strand helps explain these ATT patterns. Students described shifting from answer-hunting to planned, monitored, and flexible problem-solving, together with a

[21] ICAP stands for four cognitive engagement modes, labelled as Interactive, Constructive, Active, and Passive (Chi, 2009; Chi & Wylie, 2014).

greater willingness to articulate and defend reasoning, precisely the dispositions that higher-order items demand. Where workshop routines (modelling strategies, group problem-solving, reflective prompts) align with assessment demands (as in Test 2F's HO-dense design), these behaviours surface most clearly. This alignment is consistent with broader evidence that matching learning activities to intended higher-order outcomes, and assessing those outcomes directly is critical for detecting impact, and that active-learning approaches typically improve performance on conceptually demanding tasks in undergraduate STEM (Freeman et al., 2014; McNeill et al., 2012).

7.6.4 Qualitative Research Question 1

How do students perceive the impact of the Mathematical Thinking Workshops on their learning approaches and understanding of university mathematics?

Students' accounts (Chapter 6) converge on a clear perception: the workshops fostered reflective, strategy-rich engagement with mathematics. In the theme *Developing Reflective and Strategic Mathematical Thinking*, participants described a deliberate shift from answer-hunting to planning, monitoring, and evaluating their approaches. They reported pausing to analyse structure, classifying problems before acting, and selecting methods based on task features rather than habit. These practices reflect core principles of metacognitive awareness, thinking about and regulating one's own thinking (Flavell, 1979; Schoenfeld, 1992), and were frequently linked by students to more explicit reasoning and fewer blind procedural moves.

Across *Reconnecting with Mathematics Through Alignment with Coursework* and *Growth in Mathematical Communication and Problem-Solving Confidence*, students narrated unlearning rote routines from school and replacing them with conceptually justified steps, echoing views of meaningful learning as the coordination of concepts and procedures (Hiebert & Lefevre, 1986). They described explaining ideas aloud, listening to peers' alternative starts, and revising solutions considering feedback, practices that normalised error as information rather than deficiency. This was enabled by the *Facilitator Influence*

and Supportive Learning Environment theme, which created consistent norms (inviting questions, surfacing thinking, and asking “why this step?”) that fostered a psychological safety, making risk-taking and strategy experimentation feel legitimate.

Students also reported shifts in mathematical identity, moving from tentative participation to seeing themselves as contributors to mathematical conversations. In Sfard and Prusak’s (2005) terms, students’ narratives indicated evolving self-stories (“I’m the kind of person who can figure this out if I try different ways”), which they tied to workshop practices of modelling strategies, group problem-solving, and short reflective prompts. The qualitative evidence portrays workshops as an ecology that cultivates metacognition, strategic flexibility, and a sense of belonging, under which students perceive themselves as becoming more capable and agentic mathematical thinkers.

7.6.5 Qualitative Research Question 2

What aspects of the workshops do students identify as most beneficial for their mathematical development?

Students identified several workshop features as distinctly beneficial, and their accounts in Chapter 6 locate these benefits in how the sessions were organised, facilitated, and experienced. A central thread is alignment with coursework. In *Reconnecting with Mathematics Through Alignment with Coursework*, students described recognising problem types, representations, and solution routines from workshops in their tutorials and class tests. This perceived fit made the material feel relevant and timely, increased motivation to prepare. It helped them approach assessments with clearer expectations about how to begin and how to check work. This resonates with work on constructive alignment, which suggests that performance gains are most likely when learning activities and assessments explicitly target the intended learning outcomes (Loughlin et al., 2020; Maffei et al., 2022).

A second theme is the role of collaborative learning in catalysing understanding. In *Growth in Mathematical Communication and Problem-Solving Confidence,* students

described how small-group work prompted them to verbalise strategies, compare approaches, and refine justifications; hearing peers' alternative starts or checkpoints expanded their repertoire and normalised revision. Recent research on inquiry-based approaches in undergraduate mathematics identifies these very features, sustained engagement with rich tasks and collaboration for sense-making, as pillars of effective practice, and large-scale syntheses in STEM show that active, reasoning-oriented instruction improves performance and reduces failure rates across the achievement spectrum (Laursen & Rasmussen, 2019; Theobald et al., 2020). This aligns with a Vygotskian view of learning (ZPD), in which social interaction supports movement from tentative to more expert performance.

Third, the facilitator's role was repeatedly singled out. In *Facilitator Influence and Supportive Learning Environment*, students credited structured but inclusive facilitation—modelling "thinking aloud," posing probing "why" questions, and offering just-in-time hints, with normalising struggle, lowering anxiety, and sustaining momentum. As one student put it, the workshops "gave me more confidence… last year I felt so small and insignificant… this year I came to the workshops and then I realised everyone felt that way… I just feel I'm part of a community programme," capturing how facilitation shifted self-beliefs and belonging. The emphasis on respectful dialogue and valuing partial ideas created psychological safety, a crucial condition for students who may have felt marginalised or silenced in mathematics spaces (Downing et al., 2020; Gonda et al., 2024; Solomon, 2007).

Finally, students experienced the workshops as emotional and social support. They reported that they could be vulnerable without fear of judgement and draw energy from a community committed to shared effort. As one student put it using a rowing metaphor,

> So I do rowing, and we always have a saying: "It's not for you; it's for the crew." You're not rowing for yourself because nobody else in the boat is giving up on you, so you can't give up on yourself. So it's like I don't want to give up on my classmates. I know we're halfway through this difficult problem and we don't know

> what to do, but I don't want them to give up on me, and I'm not going to give up on them because I back you guys.

This aligns with research showing that belonging and psychological safety are central to engagement and persistence in university STEM: brief, context-sensitive belonging interventions can improve first-year progress, and climates that reduce fear of negative evaluation and foster safety enable fuller participation and learning (Brady et al., 2020; Gonda et al., 2024; Walton & Cohen, 2011; Walton et al., 2023). In students' narratives, the workshops nurtured not only cognitive growth but also belonging and agency, reinforcing the sense that they could participate meaningfully in mathematical conversations.

7.6.6 Qualitative Research Question 3

What challenges and barriers do students experience in engaging with the workshop content and implementing the learnt strategies?

While the workshops were widely experienced as helpful, participants also identified barriers that constrained engagement and follow-through. In *Student Voices on Workshop Hurdles*, students described environmental and logistical obstacles, hot, noisy, and poorly ventilated rooms; occasional overcrowding; and timetabling pressures that left some arriving late or leaving early. These conditions made it harder to concentrate, hear peers, and sustain effort on demanding problems. Consistent with these accounts, research in educational psychology confirms that physical environmental discomfort (heat, crowding, and poor ventilation) can significantly increase extraneous cognitive load, detracting from attention and learning (Choi et al., 2014). These also reports align with recent evidence that classroom heat and inadequate ventilation depress attention and learning performance, with measurable decrements in cognitive tasks as thermal load rises and air quality declines (Ragab et al., 2025; Vasilakopoulou & Santamouris, 2025).

Students also pointed to time and pacing constraints. Sessions sometimes ended before challenging tasks could be completed, and the momentum built during exploration was

lost without a chance to consolidate learning. Several participants requested clearer closure, asking for brief, structured debriefs of difficult questions and for tricky items to be revisited in subsequent sessions so that partial ideas could be refined, not abandoned. This request is well supported by current scholarship: debriefing and feedback-literacy practices, short, guided reflection and comparison against worked exemplars, enhance problem-solving, psychological safety, and transfer, and help students transform activity into learning (Nicol, 2020; Schürmann et al., 2025; Tan et al., 2024). It is worth noting that, by design, fully worked solutions to the workshop worksheets were not distributed; the pedagogical intent was to de-centre “answer-getting” and foreground explanation, strategy comparison, and revision. This design choice, while purposeful, could amplify the sense of “unfinished work” when time elapsed, hence students’ calls for short debriefs and planned revisiting of challenging items.

A further barrier involved uneven group dynamics. Not all students felt equally able to participate; confident peers occasionally dominated discussion, while quieter students hesitated to voice tentative ideas. As one student recalled, “in the beginning … you’d get into a group and no one wanted to talk, and then some other person is just yammering on.” This echoes the tension, noted alongside *Growth in Mathematical Communication and Problem-Solving Confidence*, between the benefits of collaboration and the need for facilitation moves (e.g., role rotation, turn-taking prompts) that distribute talk and ensure equitable access to feedback. Recent work has shown that participation is patterned by social identity and anxiety, and that the fear of negative evaluation can suppress contributions in active-learning settings unless instructors deliberately structure turn-taking, roles, and norms (Aguillon et al., 2020; Cooper et al., 2018). Designing for equitable talk is therefore essential if collaborative formats are to benefit all students.

Students expressed concern about the continuity of support beyond the first year. They valued the workshop space as one where struggle was normalised and reflective habits were cultivated, and several worried that these gains would fade without an ongoing structure to sustain them in later modules. In response to these concerns, adapted versions of the workshop methodologies have been introduced in selected second-year

modules. Across these accounts, the barriers are practical rather than philosophical: students endorse the workshop pedagogy but ask for learning conditions, venues, pacing, debrief routines, equitable participation structures, and longer-term continuity, that allow them to implement the strategies they are beginning to adopt.

Finally, it is important to revisit the demographic and schooling profiles of MTW participants discussed in Chapter 4. The cohort included students from a wide spectrum of school contexts, including those who had attended under-resourced public schools as well as those from more advantaged public and private schools. The group also reflected redress categories that remain closely linked to unequal historical access to educational opportunity. This combination of school-level and demographic background deepens the significance of the workshops' reported benefits, such as gains in confidence, conceptual understanding, and belonging, because these shifts occurred among students whose prior learning environments often provided uneven preparation for university mathematics. These contextual factors reinforce the conclusion that thinking-focused, metacognitive interventions like the MTWs can contribute meaningfully to narrowing transitional gaps and supporting students from diverse and historically disadvantaged educational trajectories.

7.7 Contributions to Knowledge

This study contributes to tertiary mathematics education, educational psychology, and transition support by clarifying how structured, socially engaged, and conceptually orientated workshops work, by strengthening the methods used to evaluate such interventions, and by offering actionable design guidance for scaled implementation. The contributions are theoretical, methodological, and practical in nature.

7.7.1 Theoretical Contributions

The findings provide empirical support for several foundational theories in mathematics education. The study affirms the relevance of constructivist approaches to learning, as

students demonstrated how their understanding developed through active engagement, peer discussion, and exploratory problem-solving. ZPD is practically illustrated in how students moved from dependence to autonomy through scaffolded peer and facilitator interaction within their developmental zone. The results also support APOS Theory by illustrating how students transitioned from rote actions to conceptual reasoning, treating mathematical entities as dynamic objects of thought and analysis. Furthermore, the findings expand on theories of mathematics identity formation (e.g., Sfard & Prusak, 2005; Boaler, 2002) by demonstrating that affective and identity shifts can be catalysed through community-oriented and reflective pedagogies. Students' growing confidence in writing, explaining, and defending mathematical ideas reflected both epistemological development and personal transformation.

In this way, the study bridges cognitive, social, and affective dimensions of mathematics learning, showing how they are not separate but co-constitutive aspects of meaningful engagement with the discipline.

7.7.2 Methodological Contributions

Methodologically, the work illustrates a rigorous mixed-methods causal-mechanism design for educational interventions in authentic settings. On the quantitative side, it pairs Propensity Score Matching with overlap weighting to estimate the ATT, thereby widening interpretability across the covariate space while improving balance. Transparent diagnostics, such as standardised mean differences to assess covariate balance, cluster-robust intervals, multiplicity control, effect sizes (Hedges' g, Glass's Δ), and sensitivity analysis for hidden bias, strengthen inference under plausible departures from ignorability. On the qualitative side, reflexive thematic analysis surfaces the cognitive, social, and affective mechanisms underlying observed effects. The integration step, reading quantitative patterns through qualitative mechanisms and vice versa, demonstrates how mixed methods can move beyond corroboration to explanation.

The study also models researcher reflexivity in intervention research, treating positionality as an analytic resource. Design choices, facilitation roles, and potential demand characteristics are made explicit, thereby bounding claims and aiding transfer.

7.7.3 Practical Contributions

Practically, the study provides evidence for implementing Mathematical Thinking Workshops as a complementary support structure in undergraduate mathematics programmes. It outlines a scalable model grounded in (1) collaborative problem-solving, (2) metacognitive development, (3) constructive alignment and (4) inclusive, emotionally safe facilitation.

The findings indicate that such workshops can enhance academic performance, foster strategic thinking, and promote affective resilience, while also reducing anxiety and increasing students' sense of belonging. In addition, the study offers recommendations for facilitator training and highlights the importance of interpersonal skills, motivational communication, and an understanding of student diversity. It also emphasises the need for institutional alignment, suggesting that workshops are most successful when integrated into the formal curriculum and supported by policy and infrastructure.

7.8 Recommendations

Grounded in the study’s findings, this section provides targeted recommendations for key stakeholders in tertiary mathematics education. These include recommendations for educational practitioners designing and facilitating workshops, institutional leaders and policymakers responsible for academic support, and undergraduate students engaging with mathematics. The recommendations aim to ensure that the benefits of the Mathematical Thinking Workshops can be sustained, scaled, and adapted to diverse educational contexts.

7.8.1 Recommendations for Workshop Design and Facilitation

One of the takeaways from both qualitative and quantitative evidence is the value of workshops that privilege conceptual reasoning, collaborative problem-solving, and metacognitive reflection. Students responded most strongly to tasks that required them to think aloud, justify choices, compare strategies, and reflect on what made a method appropriate. MTWs design did not replicate lectures or routine tutorials, and instead created space for exploratory, dialogic thinking, were those that most reliably supported mathematical growth and engagement.

Several design frictions limited impact, particularly later in the semester. Sessions sometimes ended without closure on demanding problems; by design, the workshops did not circulate solutions, which successfully de-centred “answer-getting” but, when time ran out, left some students unsure whether their approach was on track. A subset of activities felt under-scaffolded for students still consolidating prerequisites; they asked for clearer modelling of how to start or structure a solution without being given the final answer. Group participation was uneven in some groups, with confident peers doing most of the talking and quieter students hanging back; this diluted the intended benefit of peer explanation. Environmental constraints, hot, noisy venues and occasional overcrowding made sustained concentration difficult. Scheduling also mattered: when workshops or associated wellness activities collided with lab deadlines or peak assessment weeks, engagement dipped. Finally, gains observed on early, higher-order assessments did not consistently carry over to summative exams, signalling an alignment gap between workshop routines and high-stakes contexts.

Annotated exemplars that show setup, strategy pivots, and justification checkpoints without revealing every algebraic step can model quality without short-circuiting thinking. Pairing these with faded guidance helps: early in a topic, offer more explicit structure; as competence grows, remove supports and increase open-endedness. When groups stall, facilitators should use least-to-most hints and contingent questions so that students still do the intellectual work but are prevented from unproductive impasses.

Uneven spoken contributions can be addressed through simple facilitation routines that become part of the workshop culture: rotate roles each session; require each group to nominate a different “public voice” at the board; use brief “everyone writes” interludes before discussion to generate ideas from quieter students; and normalise error by publicly valuing partial ideas and improvement stories. A brief observation rubric (e.g., “who spoke, who wrote, who asked a question?”) can help facilitators monitor and adjust in real-time.

The attenuation of effects on final assessments suggests students need practice recruiting workshop habits under exam conditions. Build short, time-bounded “exam-style” interludes into later workshops, one complex item under mild time pressure, followed by a meta-discussion of decision points, representation switches, and checking moves. Publish weekly “approach summaries” that map workshop strategies to upcoming test constructs so that the connection to assessment is explicit without reverting to rote rehearsal.

Create a standing “revisit bank” of challenging problems that were not closed in-session. Open the following workshop with a five-minute reprise of one item from the bank, highlighting what changed in students’ approach. This practice strengthens memory, signals that struggle is respected and temporary, and prevents the frustration that students associate with unfinished work.

Consider treating facilitation as a stance, not only as content expertise. Provide training and coaching cycles on questioning, error framing, managing dominant voices, and using least-to-most scaffolds. Pair novice tutors with experienced mentors for co-facilitation and feedback. Use lightweight observation notes to support rapid, non-evaluative coaching: one thing to keep, one thing to try next time.

Protect basic learning conditions. Prioritise rooms with adequate ventilation and acoustics, cap group size to preserve talk time, and avoid scheduling workshops against lab deadlines or peak assessment clusters. Where unavoidable, switch those weeks to

consolidation-focused sessions that emphasise planning and checking rather than introducing brand-new problem types.

Track attendance and participation proxies (e.g., role rotations completed, number of groups that reached debrief) and relate these to item-level performance. This “dose–response” view enables the quick detection of groups that require additional support, topics whose scaffolding is either too light or too heavy, and sessions where closure was insufficient.

Students asked for the model to continue when the mathematics becomes more abstract. Extend the core routines into selected second-year modules, short approach-focused mini-lessons, reasoning-dense tasks, equitable participation norms, and structured debriefs, so that early gains in explanation and confidence become habits rather than isolated experiences.

These improvements keep the workshops’ core intact, reasoning-centred, collaborative, and metacognitively explicit, while addressing the specific points for improvement that students identified. The aim is not to make workshops longer or more content-heavy but to make them tighter: clearer about approach, fairer in participation, better aligned with assessment, and more resilient to the environmental and scheduling realities of a busy semester.

7.8.2 For Educational Practice

Workshops should be purposefully designed to embed metacognitive and conceptual learning goals. This means moving beyond simple procedural reinforcement and explicitly fostering reflective thinking, conceptual understanding, and strategic problem-solving. Facilitators play a crucial role by encouraging students to pause, question their assumptions, and consider multiple solution paths, helping to cultivate habits of mind that are essential for deep mathematical learning.

Collaboration and peer dialogue should be intentionally structured within these workshops. Group activities benefit from clearly defined roles, ensuring that every student can participate meaningfully and benefit from peer learning. Rather than dominating conversations, facilitators should guide discussions in ways that promote autonomy, support mutual learning, and ensure all voices are heard.

Expressive clarity in mathematical communication is another vital goal. Workshops should actively promote both verbal and written mathematical expression. Structured opportunities for students to explain their reasoning at whiteboards, write detailed justifications, and reflect on the clarity of their mathematical writing can significantly enhance both understanding and communication skills.

It is important to normalise the struggle and encourage persistence. Establishing a culture of productive failure helps students view mistakes as opportunities for learning rather than as setbacks. Facilitators can model intellectual humility by openly discussing errors and using them as teaching moments, which in turn supports students in developing resilience and confidence when facing uncertainty.

Finally, the effectiveness of workshops depends significantly on the quality of facilitation. Facilitators should receive thorough training not only in content and pedagogy but also in relational skills. Such training includes learning how to create emotionally safe environments, attune to student anxieties, and validate diverse forms of participation, thereby ensuring that all students feel supported and included throughout the learning process.

7.8.3 For Institutional Policy

Workshops should be fully integrated into the formal curriculum to maximise their impact. This can be achieved by timetabling sessions, awarding course credit, or providing other forms of official recognition. Such integration elevates the status of the workshops and

encourages consistent attendance, signalling a strong institutional commitment to students' academic and personal development.

Alignment between workshop content and course progression is crucial. Effective coordination between workshop facilitators and course lecturers ensures that workshop activities reinforce core topics and are appropriately timed to support upcoming assessments. This alignment enhances the relevance of workshop participation, enabling students to make meaningful connections between their workshop experiences and their broader coursework.

Supportive infrastructure and adequate resourcing form the backbone of successful workshop implementation. Providing well-ventilated physical spaces equipped with whiteboards and furniture suitable for group work creates an environment conducive to collaboration and active learning. Furthermore, securing funding for facilitator stipends, materials, and ongoing training is essential to maintain the quality and sustainability of the workshops.

Continuous monitoring and evaluation are necessary for ongoing improvement. Establishing precise mechanisms for collecting feedback from both students and facilitators, tracking attendance, and analysing impact data enables institutions to make informed revisions to the workshop model. This process ensures that workshops remain responsive to student needs and can adapt as new challenges or opportunities arise.

Finally, planning for sustainability and scalability is vital. Institutions should develop strategies to extend the workshop model across different faculties or campuses, considering adaptations such as peer-led sessions, hybrid delivery modes, or integration into first-year experience programmes. Thoughtful expansion should always be accompanied by efforts to maintain fidelity to the pedagogical principles that underpin the workshop's effectiveness.

7.8.4 For Future Students

Students are encouraged to engage actively and consistently by attending workshops regularly and participating fully in collaborative discussions. The most significant benefits arise when workshops are approached not as remedial sessions but as valuable opportunities for exploration, personal growth, and building a sense of community within the mathematics learning environment.

Embracing a growth mindset and engaging in reflective practice is essential. Workshops should be seen as spaces not only to strengthen mathematical skills but also to develop one's identity and mindset as a learner. Students can accelerate both cognitive and emotional growth by reflecting on their own problem-solving processes and learning from their peers.

Finally, students are advised to seek out and build support networks. This includes forming or joining peer study groups, asking questions freely, and utilising available academic and wellness support services. Cultivating such networks is vital for sustaining motivation and well-being throughout the academic journey, ensuring that students remain connected and supported as they navigate the challenges of university mathematics.

7.9 Directions for Future Research

Building on the insights and limitations of this study, several avenues for future research can deepen the understanding of how Mathematical Thinking Workshops and similar interventions support student learning. These directions encompass longitudinal, methodological, and contextual dimensions, aiming to contribute to both theoretical advancements and practical improvements in tertiary mathematics education.

7.9.1 Longitudinal Impact and Academic Trajectories

One key area for future inquiry is the long-term impact of workshop participation. This study focused on short-term academic outcomes within a single academic year. However, it did not explore whether the cognitive, affective, and identity-related gains persisted into subsequent years of study.

Future research could (a) track students over multiple years to evaluate the durability of learning strategies and identity shifts, (b) examine the influence of workshop participation on retention in mathematics-intensive programmes or broader STEM pathways, and (c) investigate whether students who participated in workshops become more likely to engage in peer support, seek help when needed, or adopt leadership roles in academic settings.

7.9.2 Workshop Design and Delivery Modes

Given the success of the intervention, future studies could explore variations in workshop design to determine which components are most impactful. For example, (a) what is the relative effect of metacognitive prompts, collaborative structures, or facilitator style? (b) How do students respond to online, hybrid, or flipped workshop models in comparison to face-to-face delivery? (c) Can the workshop model be adapted for courses with large enrolments or scaled using peer-led or AI-supported facilitation without losing effectiveness? Experimental or quasi-experimental studies comparing variations in structure and delivery could help optimise designs for different institutional contexts.

7.9.3 Affective Dimensions and Identity Transformation

While this study surfaced powerful narratives around confidence, resilience, and belonging, future research could explore the affective and identity-related outcomes of such interventions more systematically. For instance, (a) use of validated instruments to measure changes in math self-efficacy, math anxiety, and academic resilience; (b) case studies or narrative enquiries exploring how mathematical identity evolves in response to

pedagogical support. (c) comparative studies examining identity development across student demographics, including first-generation students, women in STEM, or students from under-resourced educational backgrounds. Such work would contribute to growing scholarship on the emotional and relational dimensions of mathematics education.

7.9.4 Institutional Implementation and Cross-Context Transferability

This study was conducted within a single institutional and disciplinary context. Future research could explore (a) the implementation of similar workshops in other institutions or faculties, including engineering and teacher education; (b) the barriers and enablers to institutional uptake, including funding models, policy alignment, and staff training needs; and (c) the role of institutional culture and leadership in sustaining and expanding inclusive teaching interventions. Such investigations would support more scalable and sustainable models of workshop-based learning, tailored to diverse academic settings.

7.9.5 Artificial Intelligence and the Changing Educational Landscape

Future studies should treat Artificial Intelligence (AI) as a design variable within MTWs, rather than a background threat, by experimentally comparing AI-supported workshops (e.g., critiquing AI-generated solutions or using AI to vary problems) with traditional workshops. Measure clear outcomes such as quality of reasoning, metacognitive habits, transfer to new problems, and persistence. Collect simple process evidence, drafts and version histories, short notes on how AI was used, and brief oral checks of understanding. Include process evidence, version histories, AI-use disclosures, brief oral justifications, to identify mechanisms, and test integrity-preserving assessment formats under permitted AI conditions (e.g. in-class reasoning checks).

7.9.6 Methodological Innovation and Mixed Methods Evaluation

Finally, future research should continue to push the boundaries of methodological innovation. The integration of Propensity Score Matching (and overlap-weighted models) with reflexive thematic analysis in this study provided a rich and rigorous account of

impact. Future studies might (a) extend this approach with longitudinal propensity score tracking, (b) incorporate learning analytics, classroom observations, or video analysis to complement self-reported data, and (c) employ participatory or co-design methods, involving students and facilitators in developing or evaluating interventions. This would strengthen evidence and promote student agency and shared ownership of educational change.

7.10 Final Reflections

This study aimed to investigate the impact of Mathematical Thinking Workshops on undergraduate students transitioning into university mathematics. What emerged was evidence of improved academic performance (in early first-semester assessments) and a compelling narrative of transformation in how students approach mathematics, relate to it emotionally, and come to see themselves within it.

At the heart of the workshops' success was their ability to rehumanise mathematics education. In contrast to dominant narratives that position mathematics as cold, individualistic, and intimidating, the workshops cultivated a learning environment that was collaborative, reflective, and emotionally supportive. Students learnt to pause and think, to ask questions without shame, and to engage in mathematics as a communal and meaningful pursuit. These findings challenge institutions to rethink how they support learners, particularly those who may have arrived underprepared or unsure of their place in mathematical spaces. They suggest that real growth occurs when students are intellectually stretched and emotionally held, when they are invited not only to solve problems but also to develop a confident, agentic identity as mathematical thinkers.

At a broader level, this study contributes to the ongoing call within STEM education to create more equitable and inclusive pathways. It shows what is possible when learning environments are intentionally designed for connection, clarity, and care by illuminating the voices of students, many of whom had experienced disengagement, anxiety, or marginalisation in prior mathematical experiences. The journey of this research has also

been personally transformative. As a researcher and educator, engaging deeply with students' reflections has reaffirmed the importance of listening to their struggles, their growth, and their potential. Their voices remind us that education is not merely about content delivery but about cultivating human capacity through trust, structure, and meaningful engagement.

While much work remains to scale such interventions, support diverse learners, and sustain long-term gains, this study offers a model of what can be achieved when academic rigour is coupled with an empathetic design. It stands as a testament to the power of learning spaces that are both humane and intellectual, and to the enduring impact of teaching that sees and believes in the student.

REFERENCES

Abowitz, D. A., & Toole, T. M. (2010). Mixed method research: Fundamental issues of design, validity, and reliability in construction research. *Journal of Construction Engineering and Management*, 136(1), 108-116. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000026

Adams, W.C. (2015). Conducting semi-structured interviews. In K.E.Newcomer, H.P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation* (4th ed., pp. 492-505). Jossey-Bass. https://doi.org/10.1002/9781119171386.ch19

Adhikari, R., & Timsina, T. P. (2024). An educational study focused on the application of mixed method approach as a research method. *OCEM Journal of Management, Technology & Social Sciences*, 3(1), 94–109. https://doi.org/10.3126/ocemjmtss.v3i1.62229

Adler, J., Alshwaikh, J., Essack, R., & Gcsamba, L. (2016). Mathematics education research in South Africa 2007–2015: Review and reflection. *African Journal of Research in Mathematics, Science and Technology Education, 21*(1), 1–14. https://doi.org/10.1080/18117295.2016.1265858

Aguillon, S. M., Siegmund, G. F., Petipas, R. H., Drake, A. G., Cotner, S., & Ballen, C. J. (2020). Gender Differences in Student Participation in an Active-Learning Classroom. *CBE—Life Sciences Education, 19*(2), ar12. https://doi.org/10.1187/cbe.19-03-0048

Albert, L. R., & Kim, R. (2013). Developing creativity through collaborative problem solving. *Journal of Mathematics Education at Teachers College*, *4*(2). https://doi.org/10.7916/jmetc.v4i2.627

Ali, M. S., Groenwold, R. H., Belitser, S. V., Pestman, W. R., Hoes, A. W., Roes, K. C., de Boer,A., & Klungel, O. H. (2015). Reporting of covariate selection and balance assessment in propensity score analysis is suboptimal: a systematic review. *Journal of Clinical Epidemiology*, *68*(2), 122-131. https://doi.org/10.1016/j.jclinepi.2014.08.011

Alias, M., & Sulaiman, N. L. (2017). Development of metacognition in Higher Education. Concepts and strategies. In *Metacognition and Successful Learning Strategies in Higher Education*, 22–42. IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-2218-8.ch002

Anderson, L. W., & Krathwohl, D. R. (2001). A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition. New York: Longman.

Arnon, I., Cottrill, J., Dubinsky, E. D., Oktac, A., Fuentes, S. R., Trigueros, M., & Weller, K. (2014). APOS theory: A framework for research and curriculum development in mathematics education. Springer. https://doi.org/10.1007/978-1-4614-7966-6

Asiala, M., Brown, A., De Vries, D. J., & Dubinsky, E. (1996). A framework for research and development in undergraduate mathematics education. *Research in Collegiate Mathematics Education, 2*, 1-32.

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine*, 28(25), 3083-3107. https://doi.org/10.1002/sim.3697

Austin, P. C. (2011a). An Introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Austin, P. C. (2011b). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, *10*(2), 150–161. https://doi.org/10.1002/pst.433

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine*, 34(28), 3661–3679. https://doi.org/10.1002/sim.6607

Bailey, D. H., Duncan, G. J., Cunha, F., Foorman, B. R., & Yeager, D. S. (2020). Persistence and Fade-Out of Educational-Intervention Effects: Mechanisms and Potential Solutions. *Psychological Science in the Public Interest*, *21*(2), 55-97. https://doi.org/10.1177/1529100620915848

Bandura, A. (1997). Self-efficacy: The exercise of control. W H Freeman/Times Books/ Henry Holt & Co.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544-559. https://doi.org/10.46743/2160-3715/2008.1573

Bazeley, P. (2012). Integrative analysis strategies for mixed data sources. *American Behavioral Scientist*, 56(6), 814–828. https://doi.org/10.1177/0002764211426330

Behi, R., & Nolan, M. (1996). Quasi-experimental research designs. *British Journal of Nursing*, 5 (17), 1079-81. https://doi.org/10.12968/bjon.1996.5.17.1079

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70–87. https://doi.org/10.1037/a0013247

Bhandari, P. (2022). Ethical considerations in research | Types & examples. Scribbr. Retrieved from: https://www.scribbr.com/methodology/research-ethics/

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347-364.

Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher Education Research & Development*, 18(1), 57-75.

Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1, 5-22.

Biggs, J., & Tang, C. (2011). Train-the-trainers: Implementing outcomes-based teaching and learning in Malaysian higher education. *Malaysian Journal of Learning and Instruction*, 8, 1-19.

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning*. Lawrence Erlbaum Associates.

Böhmer, B., Wills, G.(2025). COVID-19 and inequality in reading outcomes in South Africa: PIRLS 2016 and 2021. *Large-scale Assessments in Education, 13*, 24. https://doi.org/10.1186/s40536-025-00256-9

Botes, H., & Mji, A. (2010). Language diversity in the mathematics classroom: Does a learner companion make a difference? *South African Journal of Education, 30*(1): 123–138.

Borji, V., Alamolhodaei, H., & Radmehr, F. (2018). Application of the APOS-ACE Theory to improve students' graphical understanding of derivative. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(7), 2947-2967. https://doi.org/10.29333/ejmste/91451

Boughey, C., & McKenna, S. (2016). Academic literacy and the decontextualised learner. *Critical Studies in Teaching and Learning*, *4*(2), 1–9. https://doi.org/10.14426/cristal.v4i2.80

Boyatzis, R.E. (1998). *Transforming qualitative information: Thematic analysis and code development.* Thousand Oaks, CA: Sage.

Brady, S. T., Cohen, G. L., Jarvis, S. N., & Walton, G. M. (2020). A brief social-belonging intervention in college improves adult outcomes for black Americans. *Science advances*, 6(18), eaay3689. https://doi.org/10.1126/sciadv.aay3689

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper (Ed.), APA handbook of research methods in psychology (Vol. 2, pp. 57–71). Washington, DC: American Psychological Association.

Braun, V., & Clarke, V. (2014). What can ‘‘thematic analysis’’ offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-being*, *9*(1), Article 26152. https://doi.org/10.3402/qhw.v9.26152

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, *Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. London, UK: SAGE Publications.

Breen, S., & O’Shea, A. (2010). Mathematical thinking and task design. *Irish Mathematical Society Bulletin, 66*(1), 39-49. https://doi.org/10.33232/BIMS.0066.39.49

Brierley, J. A. (2017). The role of a pragmatist paradigm when adopting mixed methods in behavioural accounting research. *International Journal of Behavioural Accounting and Finance*, 6(2), 140-154.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. Kluwe (Eds.), Metacognition, motivation, and understanding (pp. 65-116). Erlbaum.

Bryman, A. (2001). *Social Research Methods*. Oxford: Oxford University Press.

Bunting, B., & Williams, D. (2017). Stories of transformation: using personal narrative to explore transformative experience among undergraduate peer mentors. Mentoring & Tutoring: *Partnership in Learning*, 25(2), 166–184. https://doi.org/10.1080/13611267.2017.1327691

Buseri, J. C. (1987). The influence of culture on pupils' questioning habits in Nigerian secondary schools. *International Journal of Science Education*, *9*(5), 579–584. https://doi.org/10.1080/0950069870090507

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Centre for Educational Assessments. (2024, October). *The National Benchmark Tests: National report, 2024 intake cycle: Data report*. University of Cape Town. https://www.nbt.ac.za/sites/default/files/NBT2024_DataReport_28102024_5.pdf

Cera, R., Mancini, M., & Antonietti, A. (2013). Relationships between Metacognition, Self-efficacy and Self-regulation in Learning. *Journal of Educational, Cultural and Psychological Studies*, *4*(7), 115-141. https://doi.org/10.7358/ecps-2013-007-cera

Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin, B. Gindis, V. S. Ageyev, & S. M. Miller (Eds.), Vygotsky's educational theory in cultural context (pp. 39–64). Cambridge University Press. https://doi.org/10.1017/CBO9780511840975.004

Chan, M. C. E. & Clarke, D. (2017). Structured affordances in the use of open-ended tasks to facilitate collaborative problem solving. ZDM Mathematics Education, 49 (6), pp.951-963. https://doi.org/10.1007/s11858-017-0876-2

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293-332.

Chamberlain, D., & Vidakovic, D. (2021). Cognitive trajectory of proof by contradiction for transition-to-proof students. *The Journal of Mathematical Behaviour*, 62, 100849. https://doi.org/10.1016/j.jmathb.2021.100849

Chang, T. H., Nguyen, T. Q., Lee, Y., Jackson, J. W., & Stuart, E. A. (2022). Flexible propensity score estimation strategies for clustered data in observational studies. *Statistics in medicine, 41*(25), 5016-5032. https://doi.org/10.1002/sim.9551

Chen, O., Retnowati, E., Chan, B. K. Y., & Kalyuga, S. (2023). The effect of worked examples on learning solution steps and knowledge transfer. *Educational Psychology*, 43(8), 914–928. https://doi.org/10.1080/01443410.2023.2273762

Chen, S., & Son, L. K. (2024). High impostors are more hesitant to ask for help. *Behavioral Sciences*, 14(9), 810. https://doi.org/10.3390/bs14090810

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73–105.

Chi, M.T.H., Adams, J., Bogusch, E.B., Bruchok, C., Kang, S., Lancaster, M., Levy, R., Li, N., McEldoon, K.L., Stump, G.S., Wylie, R., Xu, D. And Yaghmourian, D.L. (2018), Translating the ICAP Theory of Cognitive Engagement Into Practice. *Cognitive Science, 42*: 1777-1832. https://doi.org/10.1111/cogs.12626

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Choi, H. H., van Merrienboer, J. J. G., & Paas, F. (2014). Effects of the physical environment on cognitive load and learning: Towards a new model of cognitive load. *Educational Psychology Review, 26*(2), 225-244. https://doi.org/10.1007/s10648-014-9262-6

Christie, P., & Collins, C. (1982). Bantu Education: apartheid ideology or labour reproduction? *Comparative Education*, 18(1),59–75. https://doi.org/10.1080/0305006820180107

Clark, C., & Kimmons, R. (2023). Cognitive Load Theory. *EdTechnica: The Open Encyclopaedia of Educational Technology*. https://doi.org/10.59668/371.12980

Clark, M., & Lovric, M. (2009). Understanding secondary–tertiary transition in mathematics. *International Journal of Mathematical Education in Science and Technology*, *40*(6), 755–776. https://doi.org/10.1080/00207390902912878

Cochran, W., & Rubin, D. (1973). Controlling bias in observational studies: A Review. *Sankhya,3*5(4), 417-446.

Cohen, E.G., & Lotan, R.A. (2014). *Designing groupwork: Strategies for the heterogeneous classroom* (3rd ed.). Teachers College Press.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46. https://doi.org/10.1177/001316446002000104

Cooper, K. M., Downing, V. R., & Brownell, S. E. (2018). The influence of active learning practices on student anxiety in large-enrollment college science classrooms. *International journal of STEM education*, 5(1), 23. https://doi.org/10.1186/s40594-018-0123-6

Connelly, B.S., Sackett, P.R., & Waters, S.D. (2013). Balancing treatment and control groups in quasi-experiments: An introduction to propensity scoring. *Personnel Psychology, 66*(2), 407-442. https://doi.org/10.1111/peps.12020

Council on Higher Education (CHE). (2020). *Extended programmes with an integrated foundation phase: Theoretical considerations for curriculum design* (Briefly Speaking, Issue 14). Council on Higher Education. https://www.che.ac.za/file/6463/download?token=cQ65fGUq

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87-114. https://doi.org/10.1017/S0140525X01003922

Creswell, J. W. (2012*). Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.

Creswell, J.W. (2015). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Pearson Education Inc.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and Conducting Mixed Methods Research* (3rd ed.). Sage Publications.

Crossley, J. G. (2020). Assessing the non-cognitive domains: measuring what matters well. In *Learning and Performance Assessment: Concepts, Methodologies, Tools, and Applications* (pp. 1021-1046). IGI Global.

Cullinane, M. J. (2011). Helping mathematics students survive the post-calculus transition. *PRIMUS*, 21(8), 669–684. https://doi.org/10.1080/10511971003692830

Daniels, H. (2001). *Vygotsky and Pedagogy* (1st ed.). Routledge. https://doi.org/10.4324/9780203469576

De Guzmán, M., Hodgson, B. R., Robert, A., & Villani, V. (1998). Difficulties in the passage from secondary to tertiary education. *Proceedings of the International Congress of Mathematicians, Berlin, Documenta Mathematica, extra volume ICM* 1998, 747-762.

Dehejia, R., & Wahba,S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, *94*(448), 1053-1062.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0

Department of Basic Education (DBE). (2024). NSC Examinations 2023: Diagnostic Report (Book 1). Pretoria: Department of Basic Education. Available at: Department of Basic Education website (official PDF).

Department of Mathematics and Applied Mathematics, University of Cape Town. (2025). First-year mathematics course pass rates (internal records). Unpublished departmental data.

Dewe, M.B., & Coyle, A. (2014). Reflections on a study of responses to research on smoking: A pragmatic, pluralist variation on a qualitative psychological theme. *Review of Social Studies, 1*(1), 21–36.

Di Martino, P., Gregorio, F. & Iannone, P. (2023).The transition from school to university in mathematics education research: new trends and ideas from a systematic literature review. *Educational Studies in Mathematics 113*(1), 7–34 . https://doi.org/10.1007/s10649-022-10194-w

Downing, V. R., Cooper, K. M., Cala, J. M., Gin, L. E., & Brownell, S. E. (2020). Fear of Negative Evaluation and Student Anxiety in Community College Active-Learning Science Courses. *CBE—Life Sciences Education*, 19(2), ar20. https://doi.org/10.1187/cbe.19-09-0186

Du Plessis, P., & Mestry, R. (2019). Teachers for rural schools: A challenge for South Africa. *South African Journal of Education*, 39, Article e1774. https://doi.org/10.15700/saje.v39ns1a1774

du Toit-Brits, C. (2022). Exploring the importance of a sense of belonging for a sense of ownership in learning. *South African Journal of Higher Education*, 36(5), 58-76. https://doi.org/10.20853/36-5-4345

Dube, B., Nkomo, D., & Thokweng, M.A. (2024). Pragmatism: An essential philosophy for mixed methods research in education. *International Journal of Research and Innovation in Social Science, 8*(3), 1001-1010. https://doi.org/10.47772/IJRISS.2024.803073

Dubinsky, E. (1991). Constructive Aspects of Reflective Abstraction in Advanced Mathematics. In: Steffe, L.P. (eds*) Epistemological Foundations of Mathematical Experience. Recent Research in Psychology.* (pp. 160-202). Springer, New York, NY. https://doi.org/10.1007/978-1-4612-3178-3_9

Dubinsky, E., & Mcdonald, M.A. (2001). APOS: A Constructivist Theory of Learning in Undergraduate Mathematics Education Research. In: Holton, D., Artigue, M., Kirchgräber, U., Hillel, J., Niss, M., Schoenfeld, A. (eds) *The Teaching and Learning of Mathematics at University Level*. New ICMI Study Series (pp. 272-282). Springer, Dordrecht. https://doi.org/10.1007/0-306-47231-7_25

Durand-Guerrier, V., Kazima, M., Libbrecht, P., Njomang Ngansop, J., Salekhova, L., Tuktamyshov, N., & Winsløw, C. (2015). Challenges and Opportunities for Second Language Learners in Undergraduate Mathematics. In R. Barwell, P. Clarkson, A. Halai, M. Kazima, J. Moschkovich, N. Planas, M. Setati-Phakeng, P. Valero, & M Villavicencio Ubillús (Eds.), *Mathematics education and language diversity: The 21st ICMI study*. Springer, Cham. pp 85-101. https://doi.org/10.1007/978-3-319-14511-2_5

Dweck, C.S. (2006). *Mindset: The new psychology of success. How we can learn to fulfill our potential*. New York, NY: Ballantine Books.

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Ederio, N., Inocian, E., Calaca, N., & Espiritu, J.G. (2023). Ethical research practices in educational institutions: A literature review. *International Journal of Current Science Research and Review*, 06(5), 2709-2724. https://doi.org/10.47191/ijcsrr/V6-i5-02

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. https://doi.org/10.2307/2666999

Eichstaedt, K. E., Kovatch, K., & Maroof, D. A. (2013). A less conservative method to adjust for familywise error rate in neuropsychological research: the Holm's sequential Bonferroni procedure. *NeuroRehabilitation*, 32(3), 693–696. https://doi.org/10.3233/NRE-130893

Engelbrecht, J. (2010). Adding structure to the transition process to advanced mathematical activity. *International Journal of Mathematical Education in Science and Technology, 41*(2), 143-154. https://doi.org/10.1080/00207390903391890

Engelbrecht, J., & Harding, A. (2015). Interventions to improve teaching and learning in first year mathematics courses. *International Journal of Mathematical Education in Science and Technology*, 46(7), 1046-1060.

Fatqurhohman, F. (2016). Transition process of procedural to conceptual understanding in solving mathematical problems. *International Education Studies*, 9(9), 182–190. https://doi.org/10.5539/ies.v9n9p182

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health services research*, 48(6pt2), 2134-2156. https://doi.org/10.1111/1475-6773.12117

Findlay, M. (1995). Who has the right answer? Differential cultural emphasis in question/answer structures and the case of Hmong students at a Northern California High School. *Issues of Applied Linguistics*, 6, 23-38. https://doi.org/10.5070/L461005206

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fiske, E. B., & Ladd, H. F. (2004). Balancing public and private resources for basic education: school fees in post-apartheid South Africa. *Changing class: Education and social change in post-apartheid South Africa,* 57-88.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, C., & Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE—Life Sciences Education*, *6*(2), 132–139. https://doi.org/10.1187/cbe.06-09-0194

Gakhar, K.D., Kapur, V., & Kaur, N. (2010). *Propensity Score Matching Method in Quasi-Experimental Designs: An Approach to Program Evaluation of INHP-III*. (No. id: 2782)

Garraway, J., & Bozalek, V.G. (2019). Theoretical frameworks and the Extended Curriculum Programme. *Alternation - Interdisciplinary Journal for the Study of the Arts and Humanities in Southern Africa*, *26*(2), 7-35. https://doi.org/10.29086/2519-5476/2019/v26n2a2

Geduld, B. (2019). A snapshot of teachers' knowledge and teaching behaviour with regard to developing self-regulated learning. *Journal of Education (University of KwaZulu-Natal)*, (77), 60-78. https://doi.org/10.17159/2520-9868/i77a04

Gee, J. P. (2000). Identity as an analytic lens for research in education*. Review of Research in Education, 25*(1), 99-125. https://doi.org/10.3102/0091732X025001099

Geisler, S., & Rolka, K. (2021). "That wasn't the math I wanted to do!"—Students' beliefs during the transition from school to university mathematics. *International Journal of Science and Mathematics Education, 19*(3), 599-618. https://doi.org/10.1007/s10763-020-10072-y

Gierus, B., Du, T., Maduforo, A. N., Gilbert, B., & Koh, K. (2025). Prevalence and Quality of Mixed Methods Research in Educational Subdisciplines: A Systematic Review. *SAGE Open*, 15(2). https://doi.org/10.1177/21582440251335171

Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs forcing.* Mill Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* New York, NY: Aldine de Gruyter.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255–274. https://doi.org/10.3102/01623737011003255

Gonda, D., Tirpáková, A., Pavlovičová, G., & Ďuriš, V. (2024). The role of a team psychological safety feeling in teamwork in the classroom. *Heliyon*, 10(18), e37618. https://doi.org/10.1016/j.heliyon.2024.e37618

Gribbons, B. & Herman, J., (1996). True and quasi-experimental designs. *Practical Assessment, Research, and Evaluation, 5*(1), 14. doi: https://doi.org/10.7275/fs4z-nb61

Guetterman, T. C., & Fetters, M. D. (2018). Two methodological approaches to the integration of mixed methods and case study designs: A systematic review. *American Behavioral Scientist*, 62(7), 900–918. https://doi.org/10.1177/0002764218772641

Guntur, M., & Purnomo, Y. W. (2024). A meta-analysis of self-regulated learning interventions studies on learning outcomes in online and blended environments. *Online Learning*, *28*(3), 563–584. https://doi.org/10.24059/olj.v28i3.4025

Guo, S. & Fraser, W.M. (2014). *Propensity Score Analysis: Statistical Methods and Applications,* Second Edition. Thousand Oaks, CA: Sage Publications.

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science, 332*(6034), 1213–1216. https://doi.org/10.1126/science.1204820

Hardman, J (2023). A Cultural-Historical Approach towards Pedagogical Transitions: Transitions in Post-Apartheid South Africa. Bloomsbury: London

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG : an international journal of obstetrics and gynaecology, 125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Harper, D. (2011). Choosing a qualitative research method. *Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners*, 83-97. https://doi.org/10.1002/9781119973249.ch7

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator, *The Review of Economic Studies, 65(4),* 261-294. https://doi.org/10.2307/2971733

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F.K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371-404). Information Age.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 1–27). Lawrence Erlbaum Associates, Inc.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis*, 15(3), 199-236.

Hochmuth, R., Broley, L., & Nardi, E. (2021). Transitions to, across and beyond university. Research and Development in University Mathematics Education. Norwich: Routledge, pp. 191-215.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410–1412. https://doi.org/10.1126/science.1177067

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, *47*(1), 5-86. https://doi.org/10.1257/jel.47.1.5

Izzati, L., & Mahmudi, A. (2018). The influence of metacognition in mathematical problem-solving. *Journal of Physics: Conference Series*. 1097. 012107. https://doi.org/10.1088/1742-6596/1097/1/012107.

Jäder, J., Lithner, J., & Sidenvall, J. (2019). Mathematical problem solving in textbooks from twelve countries. *International Journal of Mathematical Education in Science and Technology*, *51*(7), 1120–1136. https://doi.org/10.1080/0020739X.2019.1656826

Jojo, Z.M. (2019). Mathematics Education System in South Africa. In *Education systems around the world*. IntechOpen.

Joffe, H. (2012). Thematic Analysis. In D. Harper & A. Thompson (Eds), Qualitative Research Methods in Mental Health and Psychotherapy: A Guide for Students and Practitioners (pp.209-223). Chichester, UK: Wiley-Blackwell.

Joffe, M.M., & Rosenbaum, P.R. (1999). Invited commentary: Propensity scores. *American Journal of Epidemiology, 150* (4), 327-33.

Johnson, R. B. & Christensen, L. B. (2008) Educational research: Quantitative, qualitative, and mixed approaches, 3rd ed., Sage, Thousand Oaks (US).

Johnson, C., Gitay, R., Abdel-Salam, A. G., BenSaid, A., Ismail, R., Naji Al-Tameemi, R. A., Romanowski, M. H., Kazem Al Fakih, B. M., & Al Hazaa, K. (2022). Student support in higher education: campus service utilization, impact, and challenges. *Heliyon*, *8*(12), e12559. https://doi.org/10.1016/j.heliyon.2022.e12559

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has Come. *Educational Researcher*, *33*(7), 14-26. https://doi.org/10.3102/0013189X033007014

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational psychology review*, *19*(4), 509-539. https://doi.org/10.1007/s10648-007-9054-3

Kalyuga, S. (2009). The expertise reversal effect. IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-60566-048-6.ch003

Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, *8*(9), 255. https://doi.org/10.3390/socsci8090255

Kapa, E. (2001). A metacognitive support during the process of problem solving in a computerized environment. *Educational Studies in Mathematics*, *47*(3), 317–336.

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science, 38*(5), 1008–1022. https://doi.org/10.1111/cogs.12107

Karabenick, S. & Dembo, M. (2011). Understanding facilitating self-regulated help seeking. *New Directions for Teaching and Learning*, 2011(126), 33-43. https://doi.org/10.1002/tl.442

Khoza-Shangase, K., & Kalenga, M. (2024). English additional language undergraduate students' engagement with the academic content in their curriculum in a South African speech-language and hearing training programme. *Frontiers in Education, 9*, 1-12. https://doi.org/10.3389/feduc.2024.1258358

Kiger, M.E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher*, *42*(8), 846 - 854. https://doi.org/10.1080/0142159X.2020.1755030

Kim, Y.R., & Moore, T.J. (2019). Multiple levels of metacognition: Circumstances interfering with students' spontaneous metacognitive activities. *Journal of Educational Research and Practice, 9*(1), 158-178. https://doi.org/10.5590/JERAP.2019.09.1.12

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kormos, M., Pas, S.V., & Vaart, A.V. (2023). Asymptotics of Caliper Matching Estimators for Average Treatment Effects. https://doi.org/10.48550/arXiv.2304.08373

Laursen, S. L., & Rasmussen, C. (2019). I on the prize: Inquiry approaches in undergraduate mathematics. *International Journal of Research in Undergraduate*

Mathematics Education, *5*(1), 129-146. https://doi.org/10.1007/s40753-019-00085-6

Laursen, S. L., Hassi, M., Kogan, M., & Weston, T. J. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education JRME*, *45*(4), 406-418. https://doi.org/10.5951/jresematheduc.45.4.0406

Lee, B.K., Lessler, J., & Stuart, E.A. (2011). Weight trimming and propensity score weighting. *PLOS ONE*, 6(3): e18174. https://doi.org/10.1371/journal.pone.0018174

Lee, J. (2011). Educational equity and adequacy for disadvantaged minority students: School and teacher resource gaps toward national mathematics proficiency standard. *The Journal of Educational Research*, *105*(1), 64–75. https://doi.org/10.1080/00220671.2010.519409

Lee, M.J. (2005). Treatment Effect and Sensitivity Analysis for Self-selected Treatment and Selectively Observed Response, mimeo, Singapore Management University.

Leron, U. (2003). Origins of Mathematical Thinking: A Synthesis, Proceedings CERME3, Bellaria, Italy, March, 2003.

Lewin, T., & Mawoyo, M. (2014). Student access and success: Issues and interventions in South African universities. Inyathelo: The South African Institute for Advancement.

Li, F., Morgan, K. L., & Zaslavsky, A. M. (2017). Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, *113*(521), 390–400. https://doi.org/10.1080/01621459.2016.1260466

Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic inquiry*. SAGE, Thousand Oaks, 289-331. https://doi.org/10.1016/0147-1767(85)90062-8

Lindner, R. W., & Harris, B. (1992). Self-Regulated Learning and Academic Achievement in College Students.

Lishchynska, M., Palmer, C., Lacey, S., & O'Connor, D. (2023). Is motivation the key? Factors impacting performance in first year service mathematics modules. *European Journal of Science and Mathematics Education*, *11*(1), 146-166. https://doi.org/10.30935/scimath/12529

Lo, M., & Ruef, J.L. (2020). Student or teacher? A look at how students facilitate public sensemaking during collaborative groupwork. *Journal of Urban Mathematics Education*, *13*(1), 15-33. https://doi.org/10.21423/jume-v13i1a372

Loughlin, C., Lygo-Baker, S., & Lindberg-Sand, Å. (2020). Reclaiming constructive alignment. *European Journal of Higher Education*, *11*(2), 119–136. https://doi.org/10.1080/21568235.2020.1816197

Lowe, H., & Cook, A. (2003). Mind the gap: Are students prepared for higher education? *Journal of Further and Higher Education, 27*(1), 53 - 76. https://doi.org/10.1080/03098770305629

Luneta, K. (2024) . Revisiting the shortage of effective Mathematics teachers in Southern Africa. *African Journal of Teacher Education and Development, 3*(1), a86. https://doi.org/10.4102/ajoted.v3i1.86

Mabena, N., & Mokgosi, P. N. (2021). Factors contributing to poor learner performance in mathematics: A case of selected schools in Mpumalanga Province, South Africa. *Problems of Education in the 21st Century, 79*(3), 451– 465. https://doi.org/10.33225/pec/21.79.451

Maffei, A., Boffa, E., Lupi, F., & Lanzetta, M. (2022). On the design of constructively aligned educational unit. *Education Sciences*, *12*(7), 438. https://doi.org/10.3390/educsci12070438

Majumdar, A. (2019). Thematic Analysis in Qualitative Research. In M. Gupta, M. Shaheen, & K. Reddy (Eds.), *Qualitative Techniques for Workplace Data Analysis* (pp. 197-220). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-5366-3.ch009

Mann, S. J. (2001). Alternative perspectives on the student experience: Alienation and engagement. *Studies in Higher Education*, 26(1), 7–19. https://doi.org/10.1080/03075070020030689

Masango, M. M., Van Ryneveld, L., & Graham, M. A. (2020). Barriers to the implementation of electronic textbooks in rural and township schools in South Africa. *Africa Education Review*, *17*(6), 86–118. https://doi.org/10.1080/18146627.2022.2064310

Mason, A., & Singh, C. (2010). Helping students learn effective problem-solving strategies by reflecting with peers. *American Journal of Physics*, *78*(7), 748–754. https://doi.org/10.1119/1.3319652

Mason, A. J., & Singh, C. (2016). Impact of guided reflection with peers on the development of effective problem solving strategies and physics learning. *The Physics Teacher*, *54*(5), 295–299. https://doi.org/10.1119/1.4947159

Mason, J., Burton, L., & Stacey, K. (1985). Thinking mathematically. Addison-Wesley.

McNeill M., Gosper M., & Xu J. (2012). Assessment choices to target higher order learning outcomes: the power of academic empowerment. *Research in Learning Technology*, 20. https://doi.org/10.3402/rlt.v20i0.17595

Mensah, E., & Baidoo-Anu, D. (2022). Towards quality and equitable education in South Africa: Unpacking the relationship between teacher factors, students' socioeconomic background and mathematics achievements. *Research in Mathematics*, *9*(1). https://doi.org/10.1080/27684830.2022.2088645

Mercer, N., & Littleton, K. (2007). Dialogue and the Development of Children's Thinking: A Sociocultural Approach (1st ed.). Routledge. https://doi.org/10.4324/9780203946657

Mokgwathi, M. S. (2025). Investment in Mathematics Education and Curriculum Reform in South Africa: Addressing Inequities and Bridging the Achievement Gap. In C. Lin & L. Sun (Eds.), Diversity, Equity, and Inclusion for Mathematics and Science Education: Cases and Perspectives (pp. 189-232). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-0345-1.ch006

Mokhithi, M., Campbell, A. L., Shock, J. P., & Padayachee, P. (2025). 'I call it math therapy': student narratives of growth, belonging and confidence in mathematical thinking workshops. International Journal of Mathematical Education in Science and Technology, 1–26. https://doi.org/10.1080/0020739X.2025.2564193

Morgan, D.L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1): 48–76.

Morgan, D.L. (2014) Pragmatism as a paradigm for social research. *Qualitative Inquiry 20*(8), 1045–1053. https://doi.org/10.1177/1077800413513733

Morgan, D., & Skaggs, P. (2016). Collaboration In the Zone Of Proximal Development. In *International Conference on Engineering And Product Design Education*, 8 & 9 September, Aalborg University, Denmark.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, *16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Morrison, K., & van der Werf, G. (2016). Searching for causality in educational research. Educational Research and Evaluation: *An International Journal on Theory and Practice*, 22(1-2), 1-5. https://doi.org/10.1080/13803611.2016.1195081

Moseholm, E., & Fetters, M. D. (2017). Conceptual models to guide integration during analysis in convergent mixed methods studies. *Methodological Innovations*, *10*(2), 1-11. https://doi.org/10.1177/2059799117703118

Morse, J. M. (1991). Approaches to Qualitative & Quantitative Methodological Triangulation. Nursing Research, 40, 120-123.

Mouton, M., & Rewitzky, I. M. (2025). Opportunity and outcome: A quantitative evaluation of the STEM Extended Curriculum Programme at a research-intensive university. *South African Journal of Higher Education*, *39*(2), 151–171. https://doi.org/10.20853/39-2-5927

Msomi, A., & Rzyankina, E. (2024). Bridging gaps: Enhancing holistic mathematics support in the transition from secondary school to university. *Journal of Student Affairs in Africa*, *12*(2), 51-70. https://doi.org/10.24085/jsaa.v12i2.5450

Mtshweni, B. V. (2024). Sense of belonging and academic persistence among undergraduate university students: The chain mediation effect of emotional and academic adjustment. *Journal of Psychology in Africa*, *34*(2), 176–183. https://doi.org/10.1080/14330237.2024.2335868

Mugambi, M. (2018). Linking constructivism theory to classroom practice. *International Journal of Humanities Social Sciences and Education, 5*(9), 96-104. http://doi.org/10.20431/2349-0381.0509014.

Neumann, I., Jeschke, C., & Heinze, A. (2021). First year students' resilience to cope with mathematics exercises in the university mathematics studies. *Journal für Mathematik-Didaktik, 42*(2), 307-333. https://doi.org/10.1007/s13138-020-00177-w

Nicol, D. (2020). The power of internal feedback: exploiting natural comparison processes. *Assessment & Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Njeri, M., & Taym, A. (2024). Analysing the power of socioeconomic status on access to technology-enhanced learning in secondary schools. *Research Studies in English Language Teaching and Learning*, 2(4), 223–250. https://doi.org/10.62583/rseltl.v2i4.55

Noddings, N. (2012). The caring relation in teaching. *Oxford Review of Education*, *38*(6), 771–781. https://doi.org/10.1080/03054985.2012.745047

Nowell, L.S., Norris, J.M., White, D.E., & Moules, N. (2017) Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1–13. https://doi.org/10.1177/1609406917733847

Noor, B. (2022). Students' Meta-cognition Skills and Problem-solving Strategies in Math: A Preliminary Literature Review. Global Social Sciences Review, VII(IV), 82-88. https://doi.org/10.31703/gssr.2022(VII-IV).09

Nortvedt, G. A., & Siqveland, A. (2018). Are beginning calculus and engineering students adequately prepared for higher education? An assessment of students' basic mathematical knowledge. *International Journal of Mathematical Education in Science and Technology*, 50(3), 325–343. https://doi.org/10.1080/0020739X.2018.1501826

Ntsanwisi, S. (2024). Bridging gaps in STEM education: The case for dedicated learning centres in South African townships and rural areas. *European Journal of STEM Education*, 9(1), 15. https://doi.org/10.20897/ejsteme/15481

Ogude, N., Kilfoil, W. & du Plessis, G. (2012). An institutional model for improving student retention and success at the University of Pretoria. *The International Journal of the First Year in Higher Education*, *3*(1). 21-34. https://doi.org/10.5204/intjfyhe.v3i1.97

Ogude, N.A., & Rollnick, M. (2022). Ideological positioning of Extended Curriculum Programmes – a case study of a large South African research university. *South African Journal of Higher Education, 36*(2), 222-238. https://doi.org/10.20853/36-2-4526

Opesemowo, O. A. G., & Adewuyi, H. O. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *Eurasia Journal of*

Mathematics, Science and Technology Education, 20(7), em2478. https://doi.org/10.29333/ejmste/14762

Orihara, S., Amamoto, Y., & Taguri, M. (2024). Simple and robust estimation of average treatment effects for the overlap population using model averaging. *Biostatistics & Epidemiology*, 8(1). https://doi.org/10.1080/24709360.2024.2378662

Orru, G., & Longo, L. (2019). The evolution of cognitive load theory and the measurement of its intrinsic, extraneous and germane loads: a review. In *Human Mental Workload: Models and Applications: Second International Symposium*, H-WORKLOAD 2018, Amsterdam, The Netherlands, September 20-21, 2018, Revised Selected Papers 2 (pp. 23-48). Springer International Publishing.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive Load Theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, *29*(4), 394-398. https://doi.org/10.1177/0963721420922183

Pampaka, M., Pepin, B., & Sikko, S. A. (2016). Supporting or alienating students during their transition to higher education: mathematically relevant trajectories in two educational contexts. *International Journal of Educational Research*, 79(1), 240-257. https://doi.org/10.1016/j.ijer.2016.06.008

Panko, P., Curtis, J.D., Gorrall, B.K., & Little, T.D. (2015). Alternatives to Randomized Control Trials: A Review of Three Quasi-experimental Designs for Causal Inference.

Patton, M.Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Sage.

Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative social work and Practice*, *1*(3), 261-283. https://doi.org/10.1177/1473325002001003636

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Penazzi, D., McCready-Fallon, J., & Rosser, S. (2022). Experiential and outdoor learning in the transition to university courses—The mathematics case. *PRIMUS*, *33*(2), 123–136. https://doi.org/10.1080/10511970.2022.2040665

Piaget, J. (1950). The psychology of intelligence. Harcourt, Brace.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation (pp. 451–502). Academic Press.

Pólya, G. (2014). *How to solve it: A new aspect of mathematical method* (Princeton Science Library ed.). Princeton University Press. https://doi.org/10.2307/j.ctvc773pk (Original work published 1945).

Ragab, A., Hassieb, M. M., & Mohamed, A. F. (2025). Exploring the impact of window design and ventilation strategies on air quality and thermal comfort in arid educational buildings. *Scientific reports*, *15*(1), 19596. https://doi.org/10.1038/s41598-025-01362-y

Ramaila, S., Seloane, P.M., & Mavuru, L. (2019). Extended Curriculum Programmes as a Support Mechanism to Enhance South African Undergraduate Science Students' Academic Performance in Mathematics. Education and New Developments 2019.

Ramāna, L., Veilande, I., & Zīlīte, A. (2019, May). Identifying students' ways of learning of mathematics at university level. In *society. Integration. Education. Proceedings of the International Scientific Conference* (Vol. 1, pp. 460-471).

Randolph, J.J., Falbe, K.N., Manuel, A.K., & Balloun, J.L. (2014). A step-by-step guide to propensity score matching in R. *Practical Assessment, Research and Evaluation, 19*, Article 18.

Reddy, V., Winnaar, L., Arends, F., Juan, A., Harvey, J., Hannan, S. & Isdale, K. (2022) The South African TIMSS 2019 grade 9 results: building achievement and bridging achievement gaps. Cape Town: HSRC Press. http://hdl.handle.net/20.500.11910/19286

Rhodes, M. G. (2019). Metacognition. *Teaching of Psychology*, *46*(2), 168–175. https://doi.org/10.1177/0098628319834381

Rogoff, B. (1990). Apprenticeship in thinking: Cognitive development in social context. Oxford University Press.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Ross, A. A., & Onwuegbuzie, A. J. (2012). Prevalence of mixed methods research in mathematics education. *Mathematics Educator*, *22*(1), 84-113.

Rubin, D. (1977). Assignment to Treatment Group on the Basis of a Covariate, *Journal of Educational Statistics*, *2*(1): 1-26

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3-4), 169–188. https://doi.org/10.1023/A:1020363010465

Ryan, G. W., & Bernard, H. R. (2000). Data management and analysis methods. *Handbook of qualitative research*, *2*(1), 769-802.

Ryan, A.M., Gheen, M.H., & Midgley, C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social–emotional role, and the classroom goal structure. *Journal of Educational Psychology, 90*(3), 528-535.

Rylands, L. J., & Shearman, D. (2018). Mathematics learning support and engagement in first year engineering. *International Journal of Mathematical Education in Science and Technology*, *49*(8), 1133–1147. https://doi.org/10.1080/0020739X.2018.1447699

Schlotter, M., Schwerdt, G., & Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics*, *19*(2), 109–137. https://doi.org/10.1080/09645292.2010.511821

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, *19*(4), 469–508. https://doi.org/10.1007/s10648-007-9053-4

Schoenfeld, A. H. (1985). Making sense of "out loud" problem-solving protocols. *The Journal of Mathematical Behavior*, 4(2), 171–191.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics (pp. 334–370). Macmillan Publishing Co, Inc.

Schoenfeld, A.H. (2009). Learning to Think Mathematically: Problem Solving, Metacognition, and Sense Making in Mathematics (Reprint). *Journal of Education*, 196, 1 - 38.

Schoenfeld, A.H. (2012). How we think: A theory of human decision-making, with a focus on teaching. In S. J. Cho (Ed.) The Proceedings of the 12th International Congress on Mathematical Education: Intellectual and Attitudinal Challenges (pp. 229–243). Springer.

Schöer, V., Ntuli, M., Rankin, N.A., Sebastiao, C., & Hunt, K. (2010). A blurred signal? The usefulness of National Senior Certificate (NSC) Mathematics marks as predictors of academic performance at university level. *Perspectives in Education, 28*(2), 9-18.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schürmann, V., Bodemer, D., & Marquardt, N. (2025). Applying debriefings in the context of higher education: How joint reflection fosters students' collaborative problem solving. *Social Psychology of Education*, 28(1), 62. https://doi.org/10.1007/s11218-024-09991-3

Schunk, D. H. (1984). Self-efficacy perspectives on achievement behaviour. *Educational Psychologist, 19*(1), 48-58. https://doi.org/10.1080/00461528409529281

Scott, I., Yeld, N., & Hendry, J. (2007). Higher Education Monitor: A Case for Improving Teaching and Learning in South African Higher Education. Pretoria: Council on Higher Education.

Sedibe, M. (2011). Inequality of access to resources in previously disadvantaged South African high schools. *Journal of Social Sciences*, *28*(2), 129-135.

Selden, A., & Selden, J. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*(2), 123-151.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, *34*(1), 4–36.

Setlhare, R., Wood, L. & Meyer, L. (2016). Collaborated understandings of context-specific psychosocial challenges facing South African school learners: A

participatory approach. *Educational Research for Social Change*, *5*(2), 18-34. https://doi.org/10.17159/2221-4070/2016/v5i2a2

Sfard, A., & Prusak, A. (2005). Telling identities: In search of an analytic tool for investigating learning as a culturally shaped activity. *Educational Researcher, 3*4(4), 14-22. https://doi.org/10.3102/0013189X034004014

Shannon-Baker, P. (2015). Making paradigms meaningful in mixed methods research. *Journal of Mixed Methods Research, 10*(4): 319–334. https://doi.org/10.1177/1558689815575861

Shay, S., Collier-Reed, B., Hendry, J., Marquard, S., Kefale, K., Prince, R., Steyn, S. & Mpofu-Mketwa, T. et al. (2020). From Gatekeepers to Gateways: Courses Impeding Graduation Annual Report 2019. Cape Town: University of Cape Town. http://hdl.handle.net/11427/35360.

Shu, H., & Tan, Z. (2018). Improved estimation of average treatment effects on the treated: Local efficiency, double robustness, and beyond. *arXiv* preprint *arXiv:1808.01408*. https://doi.org/10.48550/arXiv.1808.01408

Simon S. D. (2025). Experimentation without randomised controls. *Journal of Oral Rehabilitation*, 52(8), 1167–1174. https://doi.org/10.1111/joor.13960

Sindhwani, A., & Sharma, M. K. (2013). Metacognitive learning skills. *Educationia Confab*, 2(4), 68-79.

Singh, P., Teoh, S. H., Cheong, T. H., Md Rasid, N. S., Kor, L. K., & Md Nasir, N. A. (2018). The use of problem-solving heuristics approach in enhancing STEM students development of mathematical thinking. *International Electronic Journal of Mathematics Education, 13*(3), 289-303. https://doi.org/10.12973/iejme/3921

Sithaldeen, R., August, L., Phetlhu, O. & Kokolo, B. (2025), Fostering college students' sense of belonging by building help-seeking behaviors in an academic recovery program. *New Directions for Higher Education, 2025*(209), 9-14. https://doi.org/10.1002/he.20527

Slavin, R. E. (2017). Evidence-based reform in education. *Journal of Education for Students Placed at Risk, 22*(3), 178–184. https://doi.org/10.1080/10824669.2017.1334560

Solomon, Y. (2007). Not belonging? What makes a functional learner identity in undergraduate mathematics? *Studies in Higher Education*, *32*(1), 79–96. https://doi.org/10.1080/03075070601099473

Spaull, N., & Jansen, J. D. (2019). *South African schooling: The enigma of inequality. A study of the present situation and future possibilities.* Springer. https://doi.org/10.1007/978-3-030-18811-5

Stanton, J. D., Sebesta, A. J., & Dunlosky, J. (2021). Fostering Metacognition to Support Student Learning and Performance. *CBE—Life Sciences Education*, 20(2), fe3. https://doi.org/10.1187/cbe.20-12-0289

Stillman, G. (2011). Applying metacognitive knowledge and strategies in applications and modelling tasks at secondary school. In G.Kaiser, W. Blum, R. B. Ferri, & G. Stillman (Eds*.), Trends in teaching and learning of mathematical modelling* (pp. 165–180). Springer

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1–21. https://doi.org/10.1214/09-STS313

Stuart, E. A., Lee, B. K., & Leacy, F. P. (2013). Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of clinical epidemiology*, *66*(8), S84-S90.

Stuckey, H. L. (2015). The Second Step in Data Analysis: Coding Qualitative Research Data. *Journal of Social Health and Diabetes*, 3(012009), 7-10. https://doi.org/10.4103/2321-0656.140875

Stylianou, D. A., Blanton, M. L., & Rotou, O. (2015). Undergraduate students‘ understanding of proof: Relationships between proof conceptions, beliefs, and classroom experiences with learning proof. *International Journal of Research in Undergraduate Mathematics Education*, *1*(1), 91–134. https://doi.org/10.1007/s40753-015-0003-0

Sweller, J. (1988). Cognitive load during problem-solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (pp. 37–76). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Swidan, O., & Yerushalmy, M. (2016). Conceptual structure of the accumulation function in an interactive and multiple-linked representational

environment. *International Journal of Research in Undergraduate Mathematics Education*, *2*(1), 30-58. https://doi.org/10.1007/s40753-015-0020-z

Tachie, S.A. (2019). Meta-cognitive skills and strategies application: How this helps learners in mathematics Problem-solving. *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(5), em1702. https://doi.org/10.29333/ejmste/105364

Tall, D. (1991). Advanced mathematical thinking. Springer.

Tall, D. (1997). From School to University: The Transition from Elementary to Advanced Mathematical Thinking.

Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal*, 20(2), 5–24. https://doi.org/10.1007/BF03217474

Tan, K. H., Carmelie Rif, S., Mohamed Sultan, F. M., & Muslim, N. (2024). Eliciting debriefing experiences: A scoping review. *Heliyon*, 10(13), e33592. https://doi.org/10.1016/j.heliyon.2024.e33592

Tanner, K. D. (2013). Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE—Life Sciences Education*, *12*(3), 322–331. https://doi.org/10.1187/cbe.13-06-0115

Tashakkori, A., & Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. Sage Publications, Inc.

Tashakkori, A., & Teddlie, C. (2010). Sage Handbook of Mixed Methods in Social & Behavioral Research (2nd ed.). Sage Publications.

Taylor, J. A., & Mander, D. (2003). Developing study skills in a first year mathematics course. *New Zealand Journal of Mathematics*, 32(Supplementary issue), 217-225.

Teherani, A., Martimianakis, T., Senfors-Hayes, T., Wadhwa, A., & Varpio, L. (2015).Choosing a qualitative research approach. *Journal of graduate medical education, 7* (4), 669- 670.

Theobald, M. (2021). Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis. *Contemporary Educational Psychology, 66*, 101976. https://doi.org/10.1016/j.cedpsych.2021.101976

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., ... & Freeman, S. (2020). Active learning narrows achievement gaps for

underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476-6483. https://doi.org/10.1073/pnas.1916903117

Thomas, M. O. J., de Freitas Druck, I., Huillet, D., Ju, M.-K., Nardi, E., Rasmussen, C., & Xie, J. (2015). Key mathematical concepts in the transition from secondary school to university. In S. J. Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education* (pp. 265-284). Springer.

Thompson, P.W. (2020). Constructivism in Mathematics Education. In: Lerman, S. (eds) Encyclopedia of Mathematics Education. Springer. https://doi.org/10.1007/978-3-030-15789-0_31

Trigueros, M., & Oktaç, A. (2019). Task design in APOS Theory*. Avances de Investigación en Educación Matemática, 15*, 43-55. https://doi.org/10.35763/aiem.v0i15.256

Tularam, G.A. (2013). Preparedness of students in first year mathematics: Importance of placing the focus and responsibility of learning on students. *International Journal of Education, 3*(1), 24-33.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, *22*(3), 271–296. https://doi.org/10.1007/s10648-010-9127-6

Vasilaki, E., & Mavrogianni, A. (2025). Extending Cognitive Load Theory: The CLAM Framework for Biometric, Adaptive, and Ethical Learning. *Psychology International*, *7*(2), 40. https://doi.org/10.3390/psycholint7020040

Vasilakopoulou, K., & Santamouris, M. (2025). Cumulative exposure to urban heat can affect the learning capacity of students and penalize the vulnerable and low-income young population: A systematic review. PLOS Clim 4(7): e0000618. https://doi.org/10.1371/journal.pclm.0000618

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family medicine*, 37(5), 360–363.

Villabona, D., Oktaç, A., & Roa Fuentes, S. (2024). Acting on totalities of infinite processes: Constructing facets of an object conception. *ZDM Mathematics Education*, *56*(7), 1375-1387. https://doi.org/10.1007/s11858-024-01631-6

Vollstedt, M., Heinze, A., Gojdka, K., & Rach, S. (2014). Framework for examining the transformation of mathematics and mathematics learning in the transition from

school to university. In S. Rezat, M. Hattermann, & A. Peter-Koop (Eds.), Transformation: A fundamental idea of mathematics education (pp. 29–50). New York, NY: Springer.

von Glasersfeld, E. (1995). A Constructivist Approach to Teaching. In L. P. Steffe, & J. Gale (Eds.), Constructivism in Education (pp. 3-15). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vosniadou, S. (2020). Bridging secondary and higher education. The importance of self-regulated learning. *European Review, 28*(S1), S94 - S103. https://doi.org/10.1017/S1062798720000939

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* . Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* (New York, N.Y.), 331(6023), 1447–1451. https://doi.org/10.1126/science.1198364

Walton, G. M., Murphy, M. C., Logel, C., Yeager, D. S., Goyer, J. P., Brady, S. T., ... & Krol, N. (2023). Where and with whom does a brief social-belonging intervention promote progress in college?. *Science*, *380*(6644), 499-505. https://doi.org/10.1126/science.ade4420

Wan, F. (2025). Propensity Score Matching: should we use it in designing observational studies?. *BMC Medical Research Methodology, 25*(1), 25. https://doi.org/10.1186/s12874-025-02481-w

Wan, S., Bond, T. N., Lang, K., Clements, D. H., Sarama, J., & Bailey, D. H. (2021). Is intervention fadeout a scaling artefact?. *Economics of Education Review, 82*, 102090. https://doi.org/10.1016/j.econedurev.2021.102090

Weiss, H. B., Kreider, H., Mayer, E., Hencke, R., & Vaughan, M. A. (2005). Working it out: The chronicle of a mixed-methods analysis. In T. S. Weisner (Ed.), Discovering successful pathways in children's development: Mixed methods in the study of childhood and family life (pp. 47-64). Chicago, IL: University of Chicago Press

Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

White, H., & Sabarwal, S. (2014). Quasi-Experimental Design and Methods: Methodological Briefs - Impact Evaluation No. 8.

William, S.K., & Maat, S.M. (2020). Understanding students' metacognition in mathematics problem solving: A systematic review. *International Journal of Academic Research in Progressive Education and Development, 9*(3), 115-127. http://doi.org/10.6007/IJARPED/v9-i3/7847

Wolhuter, C. C. (2014). Weaknesses of South African education in the mirror image of international educational development. *South African Journal of Education, 34*(2). https://doi.org/10.15700/201412071120

Won, S., & Chang, Y. (2024). Antecedents and consequences of academic help-seeking in online STEM learning. *Frontiers in psychology*, 15, 1438299. https://doi.org/10.3389/fpsyg.2024.1438299

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Child Psychology & Psychiatry & Allied Disciplines*, 17(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Wyman, P. A., Henry, D., Knoblauch, S., & Brown, C. H. (2015). Designs for testing group-based interventions with limited numbers of social units: The dynamic wait-listed and regression point displacement designs. *Prevention Science, 16*(7), 956-966. https://doi.org/10.1007/s11121-014-0535-6

Xie, Y., Zeng, F., & Yang, Y. (2024). A meta-analysis of the relationship between metacognition and academic achievement in mathematics: From preschool to university. Acta Psychologica, 249, 104486. https://doi.org/10.1016/j.actpsy.2024.104486

Xu, Z., Zhao, Y., Zhang, B., Liew, J., & Kogut, A. (2022). A meta-analysis of the efficacy of self-regulated learning interventions on academic achievement in online and blended environments in K-12 and higher education. *Behaviour & Information Technology*, *42*(16), 2911-2931. https://doi.org/10.1080/0144929X.2022.2151935

Yin, R. K. (2014). Case Study Research: Design and Methods (5th ed.). Sage Publications.

Yin, R. K. (2018). Case Study Research and Applications: Design and Methods (6th ed.). Thousand Oaks, CA: Sage Publications.

Young, D. G., & Bunting, B. D. (2024). Rethinking college transitions: Legitimate peripheral participation as a pathway to becoming. *AERA Open, 10*. https://doi.org/10.1177/23328584241255631

Zhang, Z., Kim, H. J., Lonjon, G., & Zhu, Y. (2019). Balance diagnostics after propensity score matching. *Annals of translational medicine*, *7*(1), 16.https://doi.org/10.21037/atm.2018.12.10

Zimmerman, B. J. (1990). Self-regulating academic learning and achievement: The emergence of a social cognitive perspective. *Educational Psychology Review*, 2, 173-201.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge/Taylor & Francis Group.

Zubizarreta, J.R., Stuart, E.A., Small, D.S., & Rosenbaum, P.R. (Eds.). (2023). Handbook of Matching and Weighting Adjustments for Causal Inference (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003102670

APPENDICES

Appendix A: Ethical Approval

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Faculty of Science
University of Cape Town
Rondebosch
South Africa 7701

E-mail: melissa.densmore@uct.ac.za
Tel: 021 650-9111

Original Approval Date: 05 December 2022
Revised Date: 14 December 2024

Mr Mashudu Mokhithi
Department of Mathematics and Applied Mathematics

Enhancing the mathematical thinking practices necessary for first-year university mathematics – a case study

Dear Mr Mashudu Mokhithi, A/Prof Jonathan Shock Associate, Prof Pragashni Padayachee, A/Prof Anita Campbell

I am pleased to inform you that the Faculty of Science Research Ethics Committee has approved the above-named application for research ethics clearance, subject to the conditions listed below.

- Restrictions on involving human participants in research must be adhered to, given current concerns about the spread of Covid-19. Please ensure that you are aware of and comply with UCT policy on this, as communicated by management.
- Uphold ethical principles throughout all stages of the research, responding appropriately to unanticipated issues: please contact me if you need advice on ethical issues that arise.
- Implement the measures described in your application to ensure that the process of your research is ethically sound.
- Please take note of the expiration date of your application. To renew your application please submit an amendment along with a report on the progress of your research (activities undertaken, number of participants, any incidents).
- Please submit an amendment if you add any researchers or make substantial changes to your methodology.
- Should any incidents pertaining to research ethics arise, please report this to sci-rec@uct.ac.za.

Expiration date: 14 December 2026
Your approval code is: **FSREC 089 – 2022**

I wish you success in your research.

Yours sincerely

Signed by candidate

A/Prof Melissa Densmore
Chair: Faculty of Science Research Ethics Committee

Department of Oceanography

Professor Isabelle Ansorge
FSREC Chair 2025-2027

University of Cape Town
Private Bag Rondebosch 7701
South Africa

7th March 2025

Dear Dr Mokhithi,

Enhancing the mathematical thinking practices necessary for first-year university mathematics – a case study

I am pleased to inform you that, as the Chair of the Faculty of Science Research Ethics Committee, we have approved the amendment to your previously granted ethics clearance. The amendment involves using the grades for the entire cohorts of MAM1031/2/3/4F/S students, including tests, exams, and entry grades such as NSC and NBT, without individual consent. This approach is intended to prevent reporting on skewed data and avoid potentially misleading conclusions.

The following conditions apply:

1. Please notify students participating in the workshop that their academic performance will be compared to the overall academic performance of the class.
2. Please inform students that the study will only use a sample of students' marks, not every student's marks

Thank you, Mashudu, for your patience and understanding. The FSREC is responsible for ensuring that all research complies with ethical policies. Consequently, discussions around each application require thorough due diligence by the committee to uphold UCT's ethical standards. I am satisfied that this has been accomplished with your request to amend your application.

I wish you success in your research. Please do not hesitate to contact me should you have any queries relating to this letter.
Kind Regards,

Signed by candidate

Isabelle Ansorge
FSREC Chair

Appendix B: Participant Information and Consent Forms

DEPARTMENT OF MATHEMATICS AND APPLIED MATHEMATICS

UNIVERSITY OF CAPE TOWN
PRIVATE BAG X3
RONDEBOSCH 7701
SOUTH AFRICA

RESEARCHER: Mashudu Mokhithi
TELEPHONE: +27-21-650 5676
E-MAIL: mashudu.mokhithi@uct.ac.za
URL: http://www.mth.uct.ac.za

Informed Voluntary Consent to Participate in Research Study

Enhancing the mathematical thinking practices necessary for first-year university mathematics – a case study.

I understand that I am being invited to take part in an interview. I understand that in agreeing to participate:

1. My participation is free and voluntary.
2. The interview will last approximately 60 minutes. Notes will be written during the interview and/or audio taped. I can decline to be recorded.
3. I understand that if I feel uncomfortable in any way during the interview I have the right to decline to answer any question or to leave the interview session.
4. I understand that the researcher will not identify me by name in any reports using the information obtained from the interview.
5. Participation in this interview will not have any effect on your grade for MAM1033F. This data will be anonymized before it is published.
6. If I choose to be interviewed, I have the right to view and comment on the transcribed interview data before the findings are analyzed.
7. I have been given a copy of this consent form.
8. The outcome of this study will be shared with you via email.
9. I hereby agree / disagree (circle the applicable option) to the audio recording of my interview for this study.

I hereby agree to the audio recording of my interview for this study (tick one box). ☐ Yes ☐ No

_____	_____	_____
Name of Participant	Signature of Participant	Date
_____	_____	_____
Name of Researcher	Signature of Researcher	Date

Appendix C: Focus Group Protocols (interview question sets)

Interview Questions : Focus Group 1, 2, and 3

1. What is your overall experience attending the workshops? What do you find helpful, and what are some challenges? You may refer to your experience with regard to level of comfort, engagement with facilitators/tutors, the level of difficulty of the questions, and the language used in the workshops.

2. How have the workshops helped you improve your mathematical thinking skills? Think of an activity you have done in one of the workshops recently where you felt you learnt how to do something well. Describe the activity, what you learnt, and what might have enabled your learning.

3. In your opinion, what are the main differences between school and university mathematics? Do you feel that the workshops are adequately addressing the challenges you face in transitioning from high school mathematics to university-level mathematics? Please elaborate.

4. What do you think of the overall structure of the workshops? What aspects of the workshop design do you find most helpful? What suggestions do you have for improving the design of the workshops?

5. What was your experience like working with your group members during the workshops? How did you find the balance between group work and individual work in the workshops?

6. Do you look at the worksheets before workshops? Have you looked at any worksheets after workshops? Do you have any comments about the questions in the worksheets?

7. How well do you think the workshops prepare you for class tests?

8. If a friend were to ask whether you recommend the workshops, what would you say, and why?

Interview questions : Focus Group 4 and 5

1. **On Collaboration**
   - Can you provide specific examples where group collaboration significantly aided in your understanding of a concept?
   - How do group dynamics in these workshops differ from other group settings you've been in?

2. **On Building Prerequisite Understanding**

   - Were there specific concepts or topics that you struggled with before, but the workshops helped clarify?
   - How do these workshops reinforce foundational mathematical concepts for you?

3. **On Mathematical Logic and Problem-Solving**

   - Can you describe a specific problem or topic where the workshop shifted your approach from memorization to understanding?
   - How has your approach to tackling unfamiliar mathematical problems changed since attending these workshops?

4. **On Challenging Exercises and Motivation**

   - How do the exercises given in the workshops compare to exercises in regular tutorials and WebAssign?

- What aspects of the workshops keep you motivated, especially when confronted with a difficult problem?

5. On Overall Workshop Experience

- If you could change one thing about the workshops, what would it be?
- How do you think these workshops will impact your future studies or career?

Interview questions for Focus Group 6

- Can you describe your overall experience with the workshops? Do you have any general comments or feedback about the workshops that you'd like to share? Would you recommend these workshops to future first-year students? Why or why not?
- Can you provide an example of a mathematical concept you found particularly difficult to grasp? How did the workshop format help your understanding of the concept?
- Can you share an instance where peer interaction (during the workshop) significantly enhanced your understanding of a topic?
- How did the workshops help navigate differences in knowledge or understanding within groups?
- What aspects of the workshops do you believe have not been sufficiently explored in our discussion so far? Is there something you wish we had asked you about your experience that we have not discussed yet?

Appendix D: Sample Focus Group Transcripts

Focus Group 4 : Math Therapy – 17 August 2023 – Transcript

Interviewer

OK, great. Thank you. I was introducing myself and I am a psychologist in the in the engineering faculty. My role here is really just to facilitate a discussion around the workshops that you've been attending of Mashudu. There's no right or wrong, you can say anything you want to the more feedback and comments and thoughts and ideas that you have and that you share with us, the more useful it. For him to be able to better plan kind of for future workshops and also just to kind of amend things and change things if it's like if things aren't working. So, we're interested in what's working really well, what, how you benefited and what do you think can be shifted and changed. OK. So, feel free to speak as you wish. It is recorded but everything is anonymous. So, what we will do after this is the transcription that we'll get. We're not interested really in who said what? We're really interested in the data and the content of it. And we're interested in like are there common themes. Are there things that certain people have said that contradicts other people? So, we're just interested in kind of understanding it. So, we won't link your name at all to anything that you say and when he reports as well, he might take a quote, but we won't know who said it. And it will be really in terms of kind of a discussion around a particular topic that he's coming up with or something that's come out of the data. OK. So just are there, are there any questions around that cause unless you could be like I don't know what to say. It's not going to affect your marks at all. Not at all. Not in the least. It's not related to that at all, so. Is that? Does that sound all good? For you, yeah. OK, OK. So maybe just go around and tell me your names as a start and just tell me what you are studying. Just kind of in terms of your degrees because I know if I know from my faculty all my students have to do maths first and second year. But I'm also aware that people are doing maths in a whole lot of. Other faculties, like including law. And commerce and kind of things. So just give me a sense of who. You are and what you're studying.

[Name introductions and degree programmes]

Interviewer

OK, great. Thank you. Awesome. OK, so the first thing I want to ask you is questions around collaboration. So can you give specific examples where group collaboration significantly helped you in your understanding of a particular concept? So I have heard from previous groups that the workshops are run as in groups and so we're just curious around kind of are the examples where that has helped you understand the concept better or not? OK. Yeah, sure.

Amanda

I remember when we were doing proofs for proof by induction and proof by contradiction. The collaboration when we were collaborating honestly, it helped me because I had thought I had found a pattern in solving and in doing the proofs but there wasn't really. OK, the pattern that I had detected, it wasn't it, you know, because they're different ways of solving problems and proving them by contradiction or induction. So that really helped me. And also it has helped me ask for help more like I am less afraid to ask for help.

Naledi

I definitely agree with that. I had, I think specifically with integrals. Most recently, when we had to do the like D by DX of the integral and it was just a specific type of question I haven't seen before. I hadn't gone over and. Sorry, I also don't go to lectures. And so it was really easy to just immediately be like, oh, I don't know what this is and have someone in my group explain it to me quickly because then I don't have to, like, sit by myself and be like. Oh, I actually don't know what that is. Go through textbook or like e-mail someone the whole process and it's so nice to just be able to be like quickly. Like right now, I don't understand this.

Amanda

That’s so true, also to add on to that. I think it might take you. Maybe an hour or two to just figure out one question, whereas in a group setting it's just they just add on, they just cover the content with you within 5 minutes I'd say instead of having to go through it in that. Hour in a. Group setting. Everybody can just add on their input, you know every link to the chain is important. So everybody just adds on and creates one full chain that works perfect for everybody.

Interviewer

OK. And can you say like what enables that group? Work well. What? Do you think because it feels like you. You feel quite. Free to ask and to feel to contribute. So what do you think allows for that?

Amanda

So I think because they always tell us that one has to be the Scribbler one has to be the time keeper and they give this each one a role to do a specific task. And then afterwards you have to change. You have to rotate. So one can’t remain the Scribbler. Maybe somebody doesn't understand for instance, then when everybody rotates you get to write on the board and then you writing on the board, you then understand the concept better. So I believe your question was how does it?

Interviewer

So I'm interested in like what allows for the group to work well that you feel able to speak and talk and contribute and ask, yeah.

Amanda

Because people are good listeners at that point, when they advise you to like, OK, you can take the pen, you can write on the board. Now you feel free like everybody is just accepting and they understand what you're going through at that specific time. And that's what I think makes it easy for people.

Interviewer

OK. So you're nodding. Yeah, you agree.

Hector

Yeah, I really agree with that. But yeah, it helps because sometimes maybe you don't understand really what's going on. Maybe you're we are given a question and maybe they'll say take a pen write something. They will help you if you don't understand and write something and you end up understanding everything.

Interviewer

OK. OK, so tell me how does the group this the kind of the group work in the workshops differ from other groups that you belong to like other tutorial groups. Is there a difference?

Naledi

I mean the groups in the workshops didn't start out being perfect like in the beginning. It was really brutal, and you get in a group and no one wants to talk, and then you're the only one talking. All the other one who's not talking and then some other person is just like yammering on. And no one really knows what to do, and then we got to know each other and we got comfortable. So I feel like. That's a big difference. Like you know, tutorials, they're so big and you don't really get to. Know other people. So yeah, I think that's another part of it. It's like we have a rotating roles, but at the same time, we know each other like we spent 4 hours a week together since the start of the year.

Hector

OK. To add to that, so in the workshop, sometimes we get to do exercises. This other time we had we had to partner up with someone and then we had to be a Listener and the Other one will be a speaker for like 3 minutes and you should listen without interruption and Afterwards, when the person is done they ask questions to show that you were listening intently. So I believe those exercises also help a lot because in this collaboration that. We got we were able to and when someone is speaking, we listen and after they're done, we get to ask questions. And that interaction back and forth that the dialogue helps working together even better and with our tutorial as well. We are encouraged to collaborate with the people that are in the tutorial, but it is quite hard because I believe those exercises that we did in the workshop are quite essential because you get to we get to know each other, you know, and you get to be comfortable with speaking with one another and the tutorials. It's a whole different story. It's back to like it's back to square one. You see how we started. The workshops, it's that. No one has to speak if you're speaking. No one seems to show any interest in there. And also I think because these dialogues that we have in the workshops, they are facilitated, we have, they give us 3 minutes. You have to do it right now with the partner, whereas in tutorials you don't have to so you can work by yourself. Nobody cares if you dialogue or not. You can do your own thing, or you can partner up. Nobody just cares. But then with the workshops, especially that you have to do it at that specific time for three minutes, even if it's two minutes and you have nothing more to say, just stay in the silence and wait for the person to come up with an answer or something at a specific time.

Interviewer

So what I'm hearing is that there's something around not only the actual work around the maths, but it's around the getting to know each other. That allows for the group to be a bit more cohesive and then the learning can happen and you can ask for questions. That you need to ask, OK.

Naledi

I call it math therapy. It feels like therapy. The first sessions, especially when we all like, we sat around in a circle, everyone a bit about yourself.

Hector

We go back to grade one.

Naledi

OK. But and then you like? Take a deep breath. And yeah, like the other day, we had to talk about people in our lives who've who we're grateful for, for getting us here in a maths workshop. And then the other people who are with you are like human beings. And people that you can relate. To as opposed to just. Someone in my math class.

Amanda

That's true, because I remember that we had this one session and they said look at somebody and tell them what you said. What you're grateful for and what worked, what did not work, how did you get through your first semester? And I remember I was having this conversation with the friend of mine because they usually they just sit us Around but this time they didn't sit around, and I was with a friend of mine who's sitting right here. And in that specific session. I remember you were just cheering up. Like, no, I did this in 1st semester and it worked so much. And I'm grateful for you. And we're just Crying and now we forget that it's even a math workshop. And it goes really deeper than that.

Interviewer

Great. I like that math therapy.

Naledi

It gets really hard. I think most of us have like Math trauma from high school. Like everyone's had them maths teacher at some point tell them that they're like not good enough.

Amanda

Or Switch to Math Lit. And in those moments, your brain immediately shuts down and it shuts down You won't be able to do what you are required to do. So the workshops have really helped us like in. Being more open, if you are struggling, take time. Breathe in and out and then you will see it will get clear or go for a walk you will come back with an idea.

Interviewer

So I am hearing that there is like this educating around tools like life, life, tools around. How to be as a human being and engage with each other. That then facilitates the learning. Thank you. So I mean, this is really, really helpful to hear. So the other question I want to ask you is, were there any specific concepts or topics that you struggled with before, but the workshops helped clarify?

Hector

For me its proofs. I try to follow a proof of some question and be like I don't understand this thing, but with workshops they really help me to understand what was going on in those groups like in first semester that we had a lot of proofs, but the workshop helped us to understand those proofs.
And the proof is not only proof by contradiction but also proof by proving the theorems.

Amanda

And I think I struggled with optimization from high school, from the get-go I have been struggling with optimization. I'm like How am I supposed to visualize this thing? Because now they give you a shape. They describe a shape in words like this is how it looks like perpendicular. This is an entire shape. And then I asked them. I remember at one workshop, I'm like, how am I supposed to see this? Because now if they need draw this diagram, this 3D shape incorrectly you get the entire question. So then they gave us little 3D shapes look at the shape, understand it and identify the height, identify the width and from those components then you can now work the question right. OK, now, so let's work with the. Let's work as. A group and draw this shape and now it shows me how, even if it might be. The object itself might be slightly different, but then with the height the components are the same. Then we get the same answer, so they don't have to see the exact shape. Because I'm thinking that. OK, this has to be the exact same shape as the textbook now, instead of working on the question itself, I'm now struggling to find the shape. Now that we're starting to answer the question or approaching it rather so. I think it really helps with optimization. Just visualizing it and. Being comfortable with not having that shape, you know, just identifying the components and how to work with the components only.

Naledi

I think integrals for me. Because This is my second. Time taking the course. Because it was a different course last year, but. I hated integrals. It was just not. My thing, I think for the integral test I got 27. Percent, which is. Not my best and. I also just, yeah, I just hate it. When I saw an integral I wanted to go. Out and run a mile. I wouldn't even attempt it and now? In the workshops, we kind of have to I'm not allowed to give up because like, I'm, I'm there for taking notes. So I have to do 4 hours of integrals a week and it's like. I'm coming past that, like difficult beginning stage or I'm not too scared to do it. It's now starting. To do it and. It's really hard. It's really hard. But it's. It's like happening slowly but surely.

Interviewer

And to the workshops re-enforce, any mathematical concepts for you.

Naledi

Yeah, All of them.

Amanda

OK. Because I think without the workshops, we wouldn't, I'm thinking to myself. Yes, we do have lectures and tutorials you forced to do it. But then. Without the workshops, sometimes let's say you don't even attend the lectures at all. And You forget that maybe this is something I had to do to let me get to the workshops and you plan that we're actually doing this thing and sometimes you seeing it for the first time whereas everybody is just there knowing what they're doing, right.

So before the reinforcement you being taught there by your own peers and then some, we go back the following. Friday and seeing that we still doing. Same thing. So you just press it. One just sticks around your brain through any time. it forces you to keep up. Especially with the proofs proof by contradiction and proof by induction.

We did it and so much that in the test. We got a question that was similar to the one we did in the workshops. when we question that we did in the in the and I tell you this, it was a breeze going through it was a breeze. I was smiling through the paper because I was able to do that proof so much that I even got extra marks for good writing, so they also reinforced. A mathematical writing, good mathematical writing, and being able to read mathematics.

Simphiwe

It also provides insights in different things. I mean, going back to the group work, you do your own work individually and you think that you understand this concept really well and you'll come back and you do the group work and you would say you can either say your idea or you don't have to say your idea and your understanding of it, but there's always going to be someone there who has a different point of view and it helps you understand me because I've noticed that for me as well, especially with differentiation now going into the fundamental theorem of calculus and actually understanding that gap between. Integration and differentiation and what you're actually doing to the graph.

Interviewer

Then the other question it relates to that which I think you've touched on a little bit is how is your approach to tackling and unfamiliar mathematical problems changed? Can you describe a specific problem where the workshop shifted your approach from memorization to understanding?

Amanda

I think the first thing that came to mind when you ask that question is this one proof that we had to do when you had to do it step by step with your group. To identify line one do line 2 do line 3.

Interviewer

How has the workshop shifted your approach from memorization to understanding.

Amanda

So yes, we would memorise the proofs themselves as a group, just memorize it, sing it out loud, memorize this. But then, our understanding started shifting. When now we had to literally analyse line by line. Of the proof. So do the line. Everybody's assigned a line,

memorize it. If you want to. But then when you have to come back, you must now know how this line shifted from line 1 to line 2. How did they find line 2? So now with you memorising line one. It's you don't you can't memorize how it got to line two. So the how was important. I think that's how we shifted from memorizing to understanding.

Naledi

Yeah, that's definitely. And that's specific problem. It was the fundamental theorem. We had to every question in the group. Separate lines just be like it couldn't be consecutive lines. OK so. You had to, so it doesn't matter how well you memorize your line, you still have to understand where in the. Proof it goes. And thus understand. How the proof works, which is very sneaky.

Interviewer

And how is your approach to tackling unfamiliar maths problems changed? Since attending these workshops.

Amanda

I am less afraid. When I see a math problem. I don't get nervous. I don't panic. It's like, OK, I've never seen this before, but I go back to like the fundamental. What's the rate? I'm sorry, I just forgot a word. But like, I go back to what I know write down, OK? This is what I know and then then start saying OK fine. I can implement this, I can't implement that. And then from there the problem solves itself in certain way and it becomes easier. So yeah, I I'm less afraid when I am Approaching problems.

Simphiwe

Mine was actually when memorizing. Rather than understanding, memorizing was, I think when back in week 2. So I thought I was good at math in High school and now I think I

realized the moment I got into a university and we doing like inequalities and we're doing a bunch of just a bunch of inequalities like the basic stuff of our functions are defined and. Inequalities and trying to find the solutions that. That's when I started noticing that I have gaps between my knowledge. Like then I realized I was memorizing in high school and now I need to actually pay attention into doing my work and realizing, like, what is happening here. So it's just I think it started for me from Week 2 from that realization when [lecturer], when there was an equality side, [lecturer] asked me, So what do you think? And then I stood and I was, I was blank. Like I didn't know it. And I thought I understood the inequalities very well. And then it's like then he gave me the answer and I said. OK, from this day I can't just simply. Look at the book. Read the book. Do an example. I actually have to go back into the lines and see where the connection. And that, like reinforcement or lack of the dump of proofs, came to us.

Naledi

I think also the other thing is just. Being comfortable with being uncomfortable because now you see this question everything. They're just using mathematical symbols. And it's just strange to you, and for the first time, yes, everything is going to be uncomfortable. But then the workshops taught us that just it's OK. It's OK for the first time, let it be uncomfortable. It's OK to be uncomfortable. But then as we go into that certain lesson, everybody just eases into it. And now, now the question is more familiar. It's easier to tackle than before, so I think being OK with being uncomfortable is something that we learnt from the workshops.

Naledi

I also agree with not being afraid. I always say math can smell fear. I think it can cause I found that the only times in my life I ever really did badly in math was when I felt like I couldn't do it. Even if I studied really hard, even if I went back to it afterwards and was like, no, I definitely can do that. So to just look at it without judging yourself for not understanding it and without being afraid and without getting in your head about it like

you can be anxious about the question, but don't be anxious about being anxious about the question.

Interviewer

Some you. Know like a like an in in being in at ease with not knowing and just kind of figuring things out without the panic and anxiety around like this math issue.

Naledi

And then and all the mathematicians that have proved theorems or whatnot, they didn't just do it, it didn't just come to them, it took time, and so in the workshops has taught us to be comfortable with that. Like what she said, being comfortable with being uncomfortable. And also knowing that it's okay to not get it right the first time, just keep on going, you will eventually get there.

Interviewer

And is it fun? Yes?

Naledi

It's all fun now than it used to be. In the beginning. OK, first of all. It was like February and it's hot. There's no air conditioning, there's no ventilation. No windows. There's a lot of glass. There's not a lot of windows. And also just feeling like socially awkward, scared of everyone. And then you're going to like. The only one talking and you're like I hate everyone in this session. I want to go home. It's Friday, 3:00 PM and it's hot. I'm going to go home. And then later The maths got more understandable. Winter came. Got to know each other. People Started laughing at My jokes. But yeah, so it's like it's fun. But it. It feels like we worked hard to make it fun.

Interviewer

But it sounds a bit like the maths problems itself kind. Of strengthening at the beginning. And like your way through in the. OK, how do the exercises given in the workshops compared to exercises in regular tutorials and WebAssign? Yeah, the actual exercises that are given.

Simphiwe

Initially, talking about Week 1 to 3. The content wasn't very focused because I'm pretty sure like everyone thought like, oh, these workshops will give us more examples and then they can teach us how to do it and understand. But the first week, it was more of like, let's not teach you that. Let's teach you how to think like a mathematician. And then I personally I thought why do I need to think like a mathematician? I just need to know how. To do the mathematics. And now that I only see now that like thinking like a mathematician is way better than just like someone like telling you how to do it. This is what you need to do, and recognizing patterns and but now so on it's. It's not hard. I would say that the tutorials are harder than the workshops, but the workshops are more focused and like actually showing you different branches of how to solve it where the tutorials are just like repetitive.

Interviewer

OK. And what aspects of the workshops keep you motivated, especially when you are confronted with a difficult problem? So are the aspects of the actual workshop. That helps you with the motivation.

Naledi

Not so silly, but that they're compulsory, I'm not going to lie, it helps. If they were last year, they were optional. Did I go? No. Was that a mistake? Yes. So they like the whole 90%

attendance thing. The fact that you needed to pass the course. It I mean it just forces you. When, like you have absolutely no motivation, you don't want to do anything. You're tired Friday. It's a long week.

Simphiwe

Will agree with that yes, the DP very important.

Interviewer

OK, OK. So the fact that it is compulsory is a motivating factor. And also that there's kind of support amongst each other. Is there anything else that that kind of motivates you when you're struggling with difficult problems in the workshops? To keep you motivated.

Naledi

I think maybe that the idea of doing it for my classmates. So I do rowing and we always have a saying. It's like not for you. For the crew. Like you're not rowing for yourself because nobody else in the boat is giving up on you, so you can't give up on yourself. So it's like I don't want to give up. On my Classmates I know, we're halfway through this difficult problem and we don't know what to do, but I don't want them to give up on me and I'm not going to give up on them, Because I back, you guys.

Interviewer

OK, so there's a sense of connectivity between people. Again, if you could change one thing about the workshops, what would it be? One or more.

Simphiwe

I didn't know if they're going to stay on like this if, like, how they running, you know, workshops now, but if it's not going to, if it's going to like. If they're going to go like this. The advertisement is right in the beginning of the year, but you shouldn't like the. I remember the advertisements will like the workshops will help you and. It's just the. Way they advertise it the first time that like. They will just. There's more, more, more practice on the questions. They should more advertised as like this is a group setting where we don't give you examples. We rather help you think like a mathematician and get remembering like orientation with [Lecturer] was like he was explaining with the like MAM1000 was broken up and he said just the school so that he's augmented and you'll. Have a workshop. And then everyone, I'm pretty sure everyone thought like it was just like. So it's just extra help for math. Advertising it in a more specific way, students. Maybe more so because I know a. Lot of people might want to come.

Naledi

I like I did like that. They advertised it as mathematical thinking workshops. Although they did sort of imply that it's more useful. For math students taking other science subjects, so like physics. So yeah, it would sort of imply that it would. Be helpful for that. But and it is. But I also think that it's so I think everyone should do it like it implies that. You should you. Only really need it if you're doing physics. Which that's not true.

Interviewer

OK. So are you. Are you saying that it would be helpful for everyone to do it? For everyone to doing maths to do it.

Naledi

Except that would be a problem.

Amanda

Because space. Because how will we then collaborate? It will be so full. There won't be enough white Boards in the very space you can.

Interviewer

Anything else? Anything else that you would like to change? About the workshop. If you could change it, so I mean because I mean the other your points taken around the way in which it's advertised could be done differently. Yeah, OK. Anything else? And the venue could be changed in summer.

Naledi

To get to the point that everyone should do it. I feel like. One thing that's good about splitting the. Courses is that? People who don't. Want to do the workshops or who don't sign up to do the workshops? So people who do sign up to the workshops, they hold you to that it's compulsory, but if you don't sign up to do the workshops, you don't have to go. Because I feel like forcing the whole course to do it would just result in a room full of people and half of them. Don't want to be there. I like the way that like the course is split and just because everyone should do it. I don't think we should force everyone to do it because I mean. It works for us. For a really specific group of people. And I know a lot of other people have a lot more like. Social awkwardness, social anxiety, problems working in groups. And just certain things that make them. Work better on their own. So, I think we have a little bit of confirmation bias just cause it worked for us cause we're sociable people?

Interviewer

But I'm also hearing you saying that you feel like other people could benefit from it or want to benefit from it with, I mean, has have people expressed interest in wanting to be part of the workshops that, you know, friends of yours? Peers?

Simphiwe

For me, they wouldn't. They didn't. Expressed it but like. The are. I can go back to proofs, but when you were doing proofs they did help us with the proofs and they didn't give an example, but they did explain how to get to it and I understood why you're not going to get immediately. And also not on the second try but. That extra boost somehow, whenever I went to my tuts or when I did with my friends. I could actually. I actually like notice the gaps that they have and I like provide that knowledge where which came from the workshops. So, so. Yeah, that's what I'm saying.

Interviewer

And the last question, how do you think these workshops will impact your future studies or career?

Amanda

I advocate for collaboration. I look forward even in other subjects. You know, if we could work together. Maybe you might get it. That's how I feel. You can't. You really can't do it alone. Let's we can go to the level 5 Chris Hani, let's get the whiteboards and let's work together. That's something I'll vouch for till this day. And I think in the future, that's how I'm going to approach. And my other courses because I know and they said it won't continue till second and third year, but if they taught us how. To do it then. I can initiate it myself. Then why not have it next year? Why not have it third year so? I think collaboration. I look, I look for that.

Simphiwe

I enjoy the quotes that [the facilitator] gives. I enjoy the quotes that he gives and how in the beginning of the session at. The end of the session. Sometimes I'll be doing he likes to give life advice and like he simply said like. That we're doing integration for volume.

And he said you must try to think outside the box. It's like you. You don't only apply that too math. You apply that to everything in life. Trying to think outside of the box. So I like those quotes he gives. It blew my mind every time.

Naledi

I definitely. I think it's giving me more confidence like. I thought doing maths last year. I felt so small and insignificant, like this tiny fish in this massive sea and felt intimidated by my peers and by my lecturers. My tutors and I felt. Kind of like I can't do this, but if I can't do this, it must mean that. I don't belong here so. I'm not going to try to do this. And then to come. To the workshops and realize that everyone felt that way. Everyone couldn't do it and it's completely reasonable to not be able to do it. At first it was really like a wild moment for me of like. That just gave me so much. It just gave me so much. And confidence and understanding of people and. It just changed; it just changed my whole. It blew my mind. It changed my whole outlook on studying science. Like I came into the humanities, and I was like, I'm going to do. A little bit of math. See what happens, and then I just. Felt like this. Outside of the whole time and now I. Just feel like I'm. Yeah, yeah, part of a community program, so. We're all we're all sharing the same struggle. And win the same. Boat like last year, there was just no sense. Of that. And if you couldn't make it on your own then. You don't belong there, which is not true, And I think it's something that I'm going to be applying to my life forever. So that I don't. Have to do it on my own.

Interviewer

OK. Any other last comments about just kind of the? Impact of this? Workshop on your life just generally when your career have been heard so much from you, anything you want, anyone Add all that you have and that you sitting with that you haven't said yet.

Hector

What can I say? Said, like collaboration from the workshops. Yeah, it really helps like. It's something that, yeah, we should also. Do with other courses like to collaborate with other students like do the work that we are supposed to do.

Amanda

OK, so can I add something? I think having a balance, you know they always say that math people are strictly books. They don't have life. And this workshop really just reminded me that you can actually balance everything. In fact, you can include that life in the math. You know you can have a life with people you can. You can have a. Social life we can be. Able to interact with other people and still be good at math. You know you can have a balance and still be good at math. So I think it removes this way of saying that math people are a certain way they can't socialize, they can't work with other people. So that just removes all of it. You and if you just have to. Balance and you're OK with that because sometimes You do feel that you have a balance. Or maybe do you feel like you're investing in other parts of your life, and when you do that, your math side will fail. Seems like you letting it down, but no it's actually. It's like being OK with having a balanced life. Sometimes this feels like it's too good to be true. But the workshops like they bring it back together. Like what are you doing? You bring it back together. You can have a balanced life and everything every part. Of your life is winning. Integration and integration.

Interviewer

OK. Thank you so much. It's been really lovely. Just hearing all these comments and kind of and just seeing how rich the experience has been because it certainly feels like it's not just been about maths, it's been about, as you say, the quotes, the, the, the kind of the engagement that then the methodology of the workshop is kind of what you've spoken a lot about and that's been lovely.

Interviewer

So what I'm going to suggest is that if you have any kind of thoughts that you may have after this discussion, I'm just can feel free to drop me an e-mail or Mashudu an e-mail just to kind of saying after something that I thought about as well, that you want to just add, if there's anything that's left. But I really just want to say thank you to each one of you for coming. We've been. Struggling and thinking people to come and talk to us. For some reason I'm actually there is. Not knowing what it's about, but we. So thank you. So this is our biggest group so far. So it's been lovely to just have you come and spend your lunch time with us, and if you can encourage other people in the class to come begin to try and have it kind of in a few Thursdays and just to try and get me because one gives many voices aren't because there may be people, they maybe people are thinking. Similar to you and maybe very different to you as well. And we just want to get back the breadth of the thinking around this so that we can also plan ahead for how to do it better. And you know what to keep or not to keep and that kind of thing, yeah. So thank you. Thanks very much. OK.

Appendix E: Initial Codes

Full initial code list, and representative quotes

Initial Code	Total Number of Text Segments	Representative Quotes
Collaboration and Peer Learning	28	FG 1: "It helps you to collaborate with other people, and different people have different perspectives on a specific problem, so it helps you to learn more and grow as a student." FG4: "Everybody can just add on their input… every link to the chain is important. So everybody just adds on and creates one full chain that works…" FG4: "They always tell us that one has to be the Scribbler one has to be the time keeper…afterwards you have to change…when everybody rotates you get to write on the board…then you…understand the concept better." FG4: "It might take you…an hour or two to just figure out one question, whereas in a group setting…within 5 minutes…every link to the chain is important." FG6: "…Neil doesn't allow us to pick our own groups…we get to interact with different people and get to gain knowledge from different people…"

		FG4: “I advocate for collaboration…Let’s get the whiteboards and let’s work together.” FG2: “Working in groups…helps you to see that…someone’s method…is much simpler and more understandable than the one that you are doing.”
Metacognitive Awareness and Strategic Thinking	26	FG1: “The workshops help take away that mindset that you need to cram something to do well.” FG2: “…before writing the answer… I just look at the question… go back to my toolkit and say, OK, this is the type of method that I should apply, not just rush through the question.” FG2: “…they taught us pattern recognition…in a test or an exam…you actually recognize…OK, this question I have to approach it this way using these methods.” FG3: “..pattern recognition…when you look at a question, it automatically clicks…you can just do your thing.” FG1: “I just pause and think about the problem before I worry about the solution.” FG4: “Let’s…teach you how to think like a mathematician…recognizing patterns…”

Deep Conceptual Understanding and Unlearning Rote Habits	23	FG1: “It's really incredible because we pinned down those misconceptions that are lying there very deeply in our understanding. And we build the foundation very solidly.” FG4: “…our understanding started shifting when we had to analyse line by line… the how was important—that’s how we shifted from memorizing to understanding.” FG1: “High school math…they teach you procedure…[at university] you first need to understand the nature of the problem itself…not just…that procedure…you have crammed.” FG2: “…you can do lots of past papers, but we really have to understand the content…[they] could…integrate 2 topics…” FG5: “…in the workshops…we will be asked OK, why are we writing this here?…reinforced the understanding of the proof…without…cram[ming] it.” FG4: “Just visualizing it…identifying the components…how to work with the components only.”
Psychological Safety and Personal Growth	22	FG4: “I call it math therapy. It feels like therapy…you forget that it’s even a math workshop. And it goes really deeper than that.”

		FG6: “It was a safe space…Everyone was collaborating…if I make a mistake…they will help and they won’t judge me.” FG6: “In the workshops you are allowed to make a mistake…they’ll try…to help you…It’s the only part where I interact with people. And I’ve made lots of friends.” FG1: “…the silent agreement that nobody is going to be judgmental and everybody’s here to learn math…”
Intentional Design and Inclusive Facilitation	21	FG1: “I think the most helpful part for me is when we go as a group to the boards to solve the questions. It really helps because you get to discuss with the other students, and then you get new perspectives on the problem. That's what's helpful for me.” FG4: “They always tell us that one has to be the scribbler, one time keeper… and then you have to rotate… when everybody rotates you get to write on the board… you then understand the concept better.” FG1: “Neil encourages us to rotate markers…there’s no group leader…Everyone has a chance to speak.” FG6: “…one has to be the Scribbler…one has to be the time keeper…afterwards you have to change.”

		FG6: “Neil…[says] each group has their own pace…there’s…a meditative session…reflection…breaks…videos…also enjoy the learning.” FG3: “Maybe the tutors can…put it on the board…Make a PowerPoint…show us…how you should have approached it.” FG2: “Maybe…a document…explaining each step on how to…get to the answer.”
A Sense of Community, Belonging, and Peer Solidarity	20	FG1: “But in the workshop, it's like everyone is so involved in helping one another. If you don't understand something, nobody's judging you when you ask a question. And you can see the amount of effort that's put into preparing the workshops and stuff, it really makes you feel like there are people supporting you as opposed to before we had the workshops.” FG1: “Yeah, it feels like a community. It feels like a family when we are there.” FG6: “We’re always laughing…there’s…a lighter…mood…a great atmosphere in the workshop.” FG1: “…a sense of being supported by…group members as well as the tutors.”

		FG4: "…we know each other…we spent 4 hours a week together since the start of the year."
Confidence and Resilience in Mathematical Problem-Solving	18	FG1: "I just pause and think about the problem before I worry about the solution. So that has helped me to approach things with more confidence and be less intimidated because I've ended up believing that I can solve a problem if I have enough time to think about it and solve it." FG4: "Being comfortable with being uncomfortable… it's OK to not get it right the first time, just keep on going, you will eventually get there." FG4: "…I'm not allowed to give up…4 hours of integrals a week…It's really hard…but…it's happening slowly but surely." FG2: "I'm very confident now…I really understand the content."
Application and Relevance	17	FG2: "I think I've found them very helpful and also the topics that we do in the workshops, Most of them are the topics that we actually have covered in the MAM1033F course and I think it's really helpful to make sort of a concrete of your understanding in that topic." FG4: "…optimization…they gave us little 3D shapes…identify the height…identify the width…and…work the question."

Frustrations, Gaps, and Environmental Barriers	16	FG3: “But then you also get to work with people, and the difficulty of it is that you're working with people, and some people won't really participate. Somebody would just tell you they are just tired and that could be a bit frustrating. It's hard to work with other people because I haven't really worked with anyone in my life. It's either I’m working alone or I’m just teaching myself all this stuff. So it's a good experience to work as a group, but then it's quite challenging if you're not used to it.” FG3: “I think they should **change the venue**… book it for only us… so that it can be more private and [we can] **concentrate**… the venue is a problem.” FG1: “It’s very humid and clustered and hot in the venue…sometimes…so hot…and I can’t really focus.” FG2: “…leave out…challenging questions because of time…maybe…indicate those questions…before we start [next time]…explain…how we should approach it.” FG1: “…they’ll just want to scribble…on the board and…move on. They don’t ask you if you [are] following.” FG4: “Because space…It will be so full. There won’t be enough white boards…”

		FG4: “In the beginning…no one wants to talk…then we got to know each other and we got comfortable.”
Sense of Community	15	FG4 : “I think maybe the idea of doing it for my classmates. So I do rowing and we always have a saying. It's like not for you. For the crew. Like you're not rowing for yourself because nobody else in the boat is giving up on you, so you can't give up on yourself. So it's like I don't want to give up on my classmates. I know we're halfway through this difficult problem and we don't know what to do, but I don't want them to give up on me and I'm not going to give up on them because I back you guys.” FG6: “…in the workshop… I interact with people… I’ve made lots of friends.” FG1: “Feels like a community…a family…” FG4: “…we got comfortable…we know each other…”
Problem-Solving and Critical Thinking	14	FG 1 : “They'll bring us to the more useful systematic way of approaching problems which you first need to understand the nature of the problem itself so that you are able to do what is necessary to solve the problem, not just take out that procedure that you have crammed in your brain and attempt.”

		FG2 : "…they taught us pattern recognition… categorizing… what type of approach you will use… so in a test you recognize, 'this question I approach this way'." FG2: "This is how we should approach them…not…a memo…also the approach…" FG5: "…use what you have learned before…that understanding helps you to solve those things." FG4: "…visualizing it…identifying the components…work with the components only."
Recommendations and Positive Feedback	14	FG6: "Overall, I think it was a very great experience for me also to interact with other people because I study alone all the time. So just being in a group environment, studying with other people, that has really been eye opening for me." FG2: "I'd recommend the workshops… they give mathematical thinking that you can apply in the outside world… even in how you reason in conversations." FG1: "Everything is helpful, it's just the venue." FG2: "…more…students to join the workshops…overall pass rate…will increase." FG5: "It's a little bit more interactive…in workshops…it's…about…having your

		thinking behind it…"
Mathematical Writing and Understanding	12	FG2: "...in terms of mathematical writing because I think that it plays a major factor in terms of mathematical thinking, because sometimes the way they pose the question is because of the mathematical writing in that question, you wouldn't really understand what the question wants." FG2: "Neil also showed…at the end the final product…is much clean…there's…five bonus marks for good writing." FG4: "…I even got extra marks for good writing…reinforced…good mathematical writing, and being able to read mathematics." FG2: "…in terms of proving we didn't really prove in high school…[now] understand where does it come from?"
Physical Environment	11	FG4: "There's no air conditioning, there's no ventilation. No windows. There's a lot of glass. There's not a lot of windows." FG1: "…very humid and clustered and hot…" FG4: "…the venue could be changed in summer."

Desire for Additional Resources or Support	10	FG5: “A lot of people were really disappointed that we're not going to have the workshops next year, like for second year maths.” FG2: “Maybe… a document explaining each step… and on the following workshop… look at those questions… not necessarily give us the answer, but how we should approach [them].” FG3: “…make a PowerPoint…show us…how you should have approached it…” FG5: “…some support from the department might be helpful…hard to…organize…next year.” FG1: “I was thinking…they probably aren’t going to have this type of support for second year math.”
Facilitator Influence and Resources	10	FG1: “The way they facilitators attend to us and the way they help us realize how useful it is to be conscious about how you're learning mathematics.” FG2: “…the host Mr Neil… advising us… sharing videos… talking about math, how we should think about math… very, very helpful.”

		FG1: “Neil encourages us to rotate markers…” FG6: “…asking the people in the workshops and the tutors was…most helpful…” FG6: “Neil…[says] each group has their own pace…timer…breaks…videos…help us…enjoy the learning.”
Learning and Enjoyment	9	FG4: “It's all fun now than it used to be in the beginning.” FG6: “…we’re always laughing…there’s…a great atmosphere in the workshop.” FG1: “I like it when we start with the whole breathing exercise…calm yourself down before you do something…really useful.”
Overcoming Challenges and Gaps in Understanding	7	FG6: “Learning new techniques to solve problems and just like you thinking you know something, but actually, when you put it into practice in the workshops, when you actually have to do the exercise that you actually realize that, oh, this is the gap that I need to fill. I don't know this as much as I thought I did.” FG5: “…we will be asked why is this line like this?…reinforced the understanding of the proof…without…cram[ming].”

		FG4: “I hated integrals…now…It’s really hard…But…slowly but surely.” FG6: “…when someone explained how they did it… that’s the first time it actually made sense… from that moment onwards, I was solving it that way and I never got the question wrong.”

Appendix F: Sample Workshop Materials (Worksheet for Workshop 1A).

MAM1033F Workshop 1A 20 February

Introduction to problem-solving

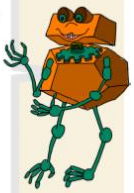
Each week we will focus on one or two mathematical thinking practices. Focus on using the relevant workshop questions to learn these practices, and then continue to use them outside of the workshops, too.

What is Mathematics?

You've probably been doing maths for a long time, but have you ever thought about what qualifies as maths and what doesn't? Since you're going to be spending a lot of time doing maths over the coming years, it's worth spending a few minutes thinking about what mathematics is.

1 On your own, complete this statement, using a language of your choice: "Mathematics is..." Share your statement with your classmates. Talk about what you agree or disagree with in these responses. Then, use the discussion to write a group statement on "Mathematics is..."

Hopefully this will make you thoroughly confused! If you're not confused, think about this: one answer you might give is that maths is the study of numbers. But lots of subjects (e.g. accounting) involve numbers and we don't refer to them as maths, while many areas of maths, like logic and algebra, don't focus on numbers. So if maths isn't simply a topic that involves numbers, then what is it?



Setting the tone for group discussions

If all went according to our cunning plan, you will have found it challenging to agree on a definition as a group. Not all the questions you'll encounter during this course will be mathematical in nature; some will involve working with people. Whether you become a researcher, go into industry, become an educator, or do something else altogether, you will likely be expected to work with others throughout your career. Figuring out the best ways to interact in social situations can be just as challenging as solving mathematical problems. We hope that you will use the problem-solving skills you learn in this course to come up with creative and effective ways of working together with other people.

1



Talking to my group about my answers slows me down. We get stuck for ages discussing one little thing.



Maybe we should stop discussing as a group to make sure we get through the workshop in time...



Okay, point taken. Here are some ideas.

That often happens when we discuss ideas, yes!



You could, but what other solutions have you considered?



2 Which of the ideas below would you consider implementing to solve Flying Sphere's (the robot on the left) problem?

- ☐ Doing the workshop individually
- ☐ Not worrying about finishing
- ☐ Not doing the workshop at all and going to the beach instead
- ☐ Setting a time limit on each question.
- ☐ Assigning a different person the job of "chair" each session, and letting that person decide when the group should move on.
- ☐ Inventing a time travelling device
- ☐ Taking longer to do the workshop
- ☐ Making notes of unresolved issues and going back to them later.
- ☐ Pretending to agree with others so that the group moves on
- ☐ Asking a tutor for the answers

3 Which of the above ideas would you consider implementing for your group?

4 Imagine you're doing group work during a workshop and one person keeps taking over. It is difficult to get a word in. You would like a chance to speak too. Which of these are likely to be constructive responses? (Select as many as you want.)

- ☐ Interrupt the person. Just keep talking if they interrupt in return.
- ☐ Say, "I'd really like a turn to say something."
- ☐ Say, "You're talking too much!"
- ☐ Go to a different group.
- ☐ Complain to a tutor
- ☐ Record the person. Play the recording back to them to let them realise how much they were talking.
- ☐ Put on an unusual hat.
- ☐ Try to exclude the person from the group next time.
- ☐ Silently fume with anger.
- ☐ Be relaxed about it; in the bigger scheme of things, this is minor.
- ☐ Have a conversation with the group about group etiquette.
- ☐ Complain about the person to your friends afterwards.

Maybe you are feeling like all this discussion about groups and interactions is a waste of time, and you'd rather just get on with the maths. That's an understandable feeling. Enjoy that feeling of being a bit uncomfortable, having to think about such a 'soft skill'. **Being comfortable with being uncomfortable is really really useful!**

In these workshops, you will often be asked to discuss your ideas with your group. Don't shy away from this. We recommend you see such discussions as part of your training as a scientist, as practice for the social interactions that are likely to be part of your job.

For a group to work together effectively, it is helpful for the group members to agree on ways to interact. We'll therefore do an exercise to come up with a set of group norms that we can all refer back to later.

5 Individually, write down (or submit online) up to five group norms/rules you would like to suggest.

We'll go through everybody's submissions, merge together similar suggestions, and publish the resulting list. Next time, everyone will be given the chance to vote for five norms they consider most important. The results of the vote will be made public once everyone has had a chance to vote.

In the meantime, here are some suggestions from us:

- **Argue**: Stand up for your ideas. Point out logical flaws when you spot them. Question assumptions and arguments.
- **Change your mind**: Practice spotting flaws in your own logic, and acknowledge when someone else has a stronger argument than you.
- **Be sensitive**: If someone gets offended by something you do or the language you use, consider what it will really cost you to change your words or behaviour. In some cases you may not understand what they're getting at, but it may cost you very little effort to show them consideration by doing what they are asking anyway.
- **Be compassionate**: Help others. Respond in a caring way if someone is having a hard day. Find a balance between robust debate and going easy on someone when they need it.

Try out your problem-solving skills

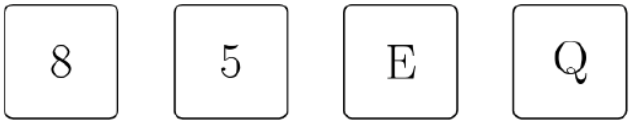
To ask and answer interesting mathematical questions, we often find it helpful to use some fairly complicated words, sentences and symbols. We will be introducing you to such complex notation and terminology later in this course, so that you will be able to engage with a large variety of interesting questions. However, the notation and terminology are not the point; they are merely tools that help us think, learn and communicate our understanding clearly. Unfortunately, when these tools are new to us, they often obscure awesome ideas instead of illuminating them. That is why we will spend this session introducing you to some interesting problems that *don't* require much prior knowledge, so that you can get a feeling for the type of thinking we hope this course will help you get a handle on.

6 A University Society has a rule for new members: when you join they flip a coin, and depending on whether it comes up heads or tails, you are told either that you must always lie, or that you must always tell the truth. If I meet three members and they make the statements below, which ones (if any) should I believe?

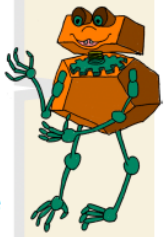
Member 1 says: "All three of us are liars."
Member 2 says: "Exactly two of us are liars."
Member 3 says: "The other two are liars."

7 Each of the four cards in the figure below has a number on one side, and a letter on the other side. Which of the card(s) should you turn over to check that the following rule is true?

If a card has an even number on one side, then it has the letter 'E' on the other side.



A cat jumps out of a window of a building that has 20 floors. It lands on solid concrete but doesn't get hurt. How did this happen? .....The window was on the ground floor!



8 **Spotting patterns** Consider a pattern made from square tiles as shown below:



a. Describe how you see this pattern growing.

b. How many tiles would you need to build the 5th figure in the pattern?

c. How many tiles would you need to build the 100th figure in the pattern?

d. Devise as many general methods as you can for predicting the number of blocks in figure n.

What is a mathematical statement?

If it is possible to assign 'true' or 'false' to a sentence/assertion, etc. then it is a mathematical statement.

Example The sentence "$x^2 - 1 = 0$" by itself is not a mathematical statement, but it becomes one when we specify a value for x. Consider the following:

(i) "For all $x \in \{1, -1\}$, $x^2 - 1 = 0$."

(ii) "For all $x \in \{1, 0\}$, $x^2 - 1 = 0$."

(iii) "There exists $x \in \{1, 0\}$ such that $x^2 - 1 = 0$."

All three are mathematical statements, the first and the third one are true, and the second one is false.

9 Which of the following are mathematical statements?

a. ☐ $3 + n + n^2$.

b. ☐ For every angle t, $\sin^2(t) + \cos^2(t) = 1$.

c. ☐ $2 \times 3 = 5$.

d. ☐ $\sqrt{2}$ is irrational.

e. ☐ $\frac{f(b) - f(a)}{b - a}$.

f. ☐ Three plus four is seven.

g. ☐ All dogs like to dance.

h. ☐ This statement is false.

i. ☐ First year lecturers know what a mathematical statement is.

j. ☐ $ax^2 + bx + c = 0$.

The quadratic equation is a tricky one! Spoiler: Think about this. Do we need to say something about all the variables for this to be a mathematical statement? For example, do we need to say something like, 'For all x...' or 'For some x...'?



k. ☐ 四等于五减一

Here's something to think about: When you're faced with some mathematical terms you don't know, can you comment on whether you're dealing with a mathematical statement or not?



Appendix G: Higher Order Questions

Table of included and excluded questions.

Assessment	Included	Excluded
Test 1F	C4, C6, C7	C1, C2, C3, C5
Test 2F	C1, C2, C4, C5, C6, C7	C3
Exam F	C1, C4, C6	C2, C3, C5
Test 1S	C3, C4, C5	C1, C2, C6
Test 2S	C1, C2	C3, C4, C5, C6
Exam S	C1, C2, C4	C3, C5, C6, C7, C8, C9, C10

TEST 1F — Section C: Long Answer Questions

Question C1. [5 points] Solve for x if $\frac{3}{x} + x \geq 4$. Write your final answer in set notation.

Question C2. [6 points] Let $f(x) = \dfrac{3x-4}{2x+1}$.

(a) Prove that f is one-to-one.

(b) Determine f^{-1}.

Question C3. [5 points] Use transformations and sketch $y = \dfrac{1}{-x+5} + 4$. Make sure that your final sketch includes intercepts.

Question C4. [4 points] If $f : \mathbb{R} \to \mathbb{R}$ is an odd function and $g : \mathbb{R} \to \mathbb{R}$ is an even function, prove that $h(x) = x^2 f(x) + x^3 g(x)$ is an odd function.

Question C5. [3 points] Consider $f(x) = x + 5$ and $h(x) = 4x - 5$. Determine g such that $g \circ f = h$.

Question C6. [4 points] Let $f(x) = |2x - 4| + |4 - 4x|$. Write f as a piecewise function and sketch it.

Question C7. [4 points] True or False? If it is true, prove it. If it is false, provide a counterexample.

(a) The composition of two decreasing functions is increasing.

(b) Given two functions f and g, $D_{f \circ g} = D_{g \circ f}$.

TEST 2F — Section C: Long Answer Questions

Question C1. [5 points] Use proof by induction to prove that $n^3 - n$ is divisible by 6 for all $n \in \mathbb{N}$.

Question C2. [5 points] Prove that $\log_{10} 3$ is irrational.

Question C3. [5 points] Assuming the limits $\lim_{h\to 0} \dfrac{\sin h}{h} = 1$ and $\lim_{h\to 0} \dfrac{\cos h - 1}{h} = 0$, prove, using the limit definition of the derivative, that if $f(x) = \cos x$ then $f'(x) = -\sin x$. Indicate all the limit laws that you used.

Question C4. [4 points] Use the Squeeze Theorem to prove that $\lim\limits_{x\to 0} x^4 e^{\cos(1/x)} = 0$.

Question C5. [4 points] Prove that the equation $9^x = 2 - x$ has a solution in the interval $[0, 1]$.

Question C6. [5 points] Consider the function $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} 3px^2 + 4, & x \geq 1, \\ p + 2x, & x < 1\ . \end{cases}$$

(a) Find the real value p such that f is continuous everywhere. (2)

(b) Using the limit definition of the derivative, prove that f is *not* differentiable everywhere. (3)

Question C7. [4 points] Using the limit definition of the derivative, prove that if $f : \mathbb{R} \to \mathbb{R}$ is an odd differentiable function, then f' is an even function.

EXAM F — Section C: Long Answer Questions

Question C1. [15 points] Consider the function $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = \dfrac{2-x}{e^x}$.

(a) Calculate the coordinates of the x- and y-intercepts. (2)

(b) Calculate the horizontal asymptotes of f. (4)

(c) Given that $f'(x) = e^{-x}(x-3)$, determine where f is increasing/decreasing by using a sign line or sign table. (2)

(d) Find the coordinates of the local extrema and classify them. (2)

(e) Given that $f''(x) = e^{-x}(4-x)$, determine the intervals where the function f is concave up/concave down. Also write down the coordinates of the point of inflection. (3)

(f) Graph the function $y = f(x)$ by showing the features found in (a)–(e). (2)

Question C2. [5 points] + 1 bonus mark for good mathematical writing. Find the point on the curve $y = \sqrt{x}$ that is closest to the point $P(3, 0)$. Fully justify your answer with calculus.

Question C3. [5 points] + 1 bonus mark for good mathematical writing. Suppose $f : \mathbb{R} \to \mathbb{R}$. Prove either part (a) or part (b) (but not both!):

(a) If $f'(x) > 0$ for all $x \in (a, b)$, then f is increasing on (a, b).

(b) If f is differentiable at $x = c$, then f is continuous at $x = c$.

Question C4. [6 points] + 1 bonus mark for good mathematical writing. Find the area under the curve $y = 6x - 3x^2$ for $0 \le x \le 2$ using a Riemann sum. Ensure to use right endpoints. *Hint:* You will need the sum in question C6 and $\sum_{i=1}^{n} i = \dfrac{n(n+1)}{2}$.

Question C5. [9 points] + 1 bonus mark for good mathematical writing.

(a) Show that the equation $x^3 + 3x = 15$ has at least one real solution. (3)

(b) Show that the equation $x^3 + 3x = 15$ has at most one real solution. (3)

(c) Thus, by (a) and (b), $x^3 + 3x = 15$ has exactly one root. Use Newton's method with initial approximation $x_1 = 2$ to find the next approximation x_2 of this root. (3)

Question C6. [5 points] + 1 bonus mark for good mathematical writing. Use proof by induction to prove that

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

for all positive integers n.

TEST 1S — Section C: Long Answer Questions

Question C1. [7 points] Evaluate $\int (x^2+1)e^{-x}\,dx$.

Question C2. [6 points] Consider the indefinite integral

$$\int \frac{(3x-1)}{x(x-1)}\,dx.$$

(a) Decompose the integrand using partial fraction decomposition and solve for all the unknown coefficients. (3)

(b) Use your answer in (a) to evaluate the integral. (3)

Question C3. [*points as per paper*] The region R is bounded by the curves $y = x+2$ and $y = 1+x^2$.

(a) Sketch the curves on the same set of axes, shade the region R and give the coordinates of the points of intersection of the two curves. (2)

(b)(i) Write down, but **do not** evaluate, the definite integral(s) that would give the volume of the solid generated by rotating the region R around the line $y = 4$. Use the disks/washers method. (4)

(b)(ii) Write down, but **do not** evaluate, the definite integral(s) that would give the volume of the solid generated by rotating the region R around the line $x = -1$. Use the cylindrical shells method. (4)

Question C4. [6 points] Use the comparison theorem to decide whether the improper integral

$$\int_6^\infty \frac{x^2+1}{x^3(\cos^2 x+1)}\,dx$$

diverges or converges.

Question C5. [4 points] Prove that if f is a continuous even function on $[-a,a]$, then for any positive real number a, $\int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx$.

Question C6. [8 points]

(a) Prove using integration by parts that if $I_n = \int \frac{dx}{(a+x^2)^n}$ for $a \neq 0$, then

$$I_{n+1} = \frac{(2n-1)I_n}{2na} + \frac{x}{(a+x^2)^n}.$$

(5)

(b) Hence calculate $\int \frac{dx}{(3+x^2)^3}$. (3)

TEST 2S — Section C: Long Answer Questions

Question C1. [7 points]

(a) Sketch the region $S_1 = \{z \in \mathbb{C} \mid 1 < |z| \leq 2,\ -\pi/4 < \mathrm{Arg}(z) \leq \pi/4\}$. (3)

(b) Now sketch the region that you obtain by applying the function $f(z) = 2iz$ to S_1. (4)

Question C2. [10 points]

(a) Write down a differential equation describing a function y whose second derivative is two times its first derivative minus two times the value of the function. (3)

(b) Find a general solution to this differential equation. (4)

(c) Fix the constants of integration by finding a solution which satisfies $y(0) = 1$ and $y'(0) = 0$. (3)

Question C3. [6 points] Determine all $z \in \mathbb{C}$ for which $\cos z = 2i$.

Question C4. [8 points] In this question you will prove De Moivre's Theorem: for any integer n and real number θ,

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta).$$

(a) First, prove the result for $n \geq 0$. (5)

(b) Now, prove the result for $n < 0$. (3)

Question C5. [4 points] Calculate the limit $\lim_{x\to 0} \dfrac{\cos 2x - 1 + 2x^2}{x^4}$ using polynomial approximations.

Question C6. [4 points] Calculate all of the 4th complex roots of $1 + \sqrt{3}i$ and plot them in the complex plane.

EXAM S — Section C: Long Answer Questions

Question C1. [8 points]

(a) Sketch the region $S_1 = \{z \in \mathbb{C} \mid -1 < \Re(z) \leq 1,\ 0 < \Im(z) \leq \pi/4\}$. (3)

(b) Now sketch the region that you obtain by applying the function $f(z) = 2e^z$ to S_1. (5)

Question C2. [11 points] The region R is bounded by the y-axis and the curves $y = \sqrt{x}$ and $y = x - 2$.

(a) Sketch the curves on the same set of axes, shade the region R and give the coordinates of the points of intersection of the two curves. If you cannot get the point of intersection, just label it as $x = a$. (3)

(b)(i) Write down, but **do not** evaluate, the definite integrals that would give the volume of the solid generated by rotating the region R around the line $y = 3$. Use the disks/washers method. (4)

(b)(ii) Write down, but **do not** evaluate, the definite integrals that would give the volume of the solid generated by rotating the region R around the line $x = -1$. Use the cylindrical shells method. (4)

Question C3. [6 points]

(a) Find a general solution to the differential equation $x^2y' + 2xy = \sin x$. (4)

(b) Fix the constant of integration by finding a solution which satisfies $y(\pi) = 1$. (2)

Question C4. [4 points] Prove the reduction formula $\int (\ln x)^n dx = x(\ln x)^n - n\int (\ln x)^{n-1} dx$.

Question C5. [3 points] Determine all $z \in \mathbb{C}$ for which $z^2 + \overline{z}^2 = 0$.

Question C6. [4 points] Prove that if y_1 and y_2 are solutions to a second-order linear homogeneous differential equation with constant coefficients, then so is $y = k_1y_1 + k_2y_2$ for any constants k_1, k_2.

Question C7. [4 points] Calculate the limit $\lim\limits_{x\to 1} \dfrac{\sin(1) - \sin x}{x - 1}$ using a Taylor approximation.

Question C8. [5 points] Calculate $\displaystyle\int_1^\infty \frac{1}{\sqrt{x-1}}\,dx$.

Question C9. [2 points] Prove that if n and k are integers with $0 \leq k \leq n$, then $\binom{n}{k} = \binom{n}{n-k}$.

Question C10. [4 points]

(a) Write down the symmetric equations of a line that passes through the point $(1, 2, 3)$ and is parallel to the vector $\langle 2, 1, 3\rangle$. (2)

(b) Calculate the intersection of the line with the (x, y)-plane. (2)

Appendix H: Python Codes for Quantitative Analysis

This appendix contains annotated Python codes that reproduces the quantitative analyses reported in Chapter 5 of the thesis (PSM estimates, overlap weighting, covariate balance diagnostics, outcome models, multiple-testing adjustment, subgroup analyses, and higher-order question analyses).

H.1 Propensity Score Estimation

```
: import pandas as pd
  import numpy as np
  from sklearn.linear_model import LogisticRegression
  from sklearn.preprocessing import StandardScaler

  # --- paths ---
  INPUT_PATH = "University_Math_Scores_4_covariates.xlsx"   # adjust if needed
  OUTPUT_PATH = "propensity_scores.xlsx"

  # --- column names ---
  ID_COL = "Student ID"
  TREAT_COL = "Augmented"
  COVARIATE_COLS = [
      "Math Grd12 Fin Rslt",
      "Eng Grd12 Fin Rslt",
      "NBT QL Score",
      "NBT Math Score",
  ]

  # 1) load
  df = pd.read_excel(INPUT_PATH, sheet_name=0)
  df.columns = [c.strip() for c in df.columns]  # tidy spaces

  # 2) build X and y
  X = df[COVARIATE_COLS].copy()
  y = df[TREAT_COL].astype(int).values

  # 3) numeric impute + cast
  for c in X.columns:
      X[c] = X[c].astype(float).fillna(X[c].median())

  # 4) standardize for stability
  scaler = StandardScaler()
  X_scaled = scaler.fit_transform(X.values)

  # 5) fit logit and compute propensity scores
  logit = LogisticRegression(max_iter=1000, solver="lbfgs")
```

```
logit.fit(X_scaled, y)
pscore = np.clip(logit.predict_proba(X_scaled)[:, 1], 1e-6, 1-1e-6)

# 6) save
df_out = df[[ID_COL, TREAT_COL] + COVARIATE_COLS].copy()
df_out["pscore"] = pscore
df_out.to_excel(OUTPUT_PATH, index=False)

print("Saved:", OUTPUT_PATH)
print("pscore summary:", {
    "min": float(df_out["pscore"].min()),
    "mean": float(df_out["pscore"].mean()),
    "max": float(df_out["pscore"].max()),
})
```

```
Saved: propensity_scores.xlsx
pscore summary: {'min': 0.0034061733808829577, 'mean': 0.2039472360095676,
'max': 0.9101780199977276}
```

H.2 Descriptive statistics of Cohort A and Cohort B

```
import pandas as pd, numpy as np, os

INPUT_PATH = "University_Math_Scores_4_covariates.xlsx"
OUT_PATH = "table4_descriptives_cohorts.xlsx"

df = pd.read_excel(INPUT_PATH)
df.columns = [c.strip() for c in df.columns]


# df["Cohort_Label"] = df["cohort"]
df["Cohort_Label"] = df["Augmented"].map({1: "Cohort A", 0: "Cohort B"}).fillna("Unknown")

VARS = ["Math Grd12 Fin Rslt","Eng Grd12 Fin Rslt","NBT QL Score","NBT Math Score"]

def summarize_by_group(data, group_col, vars_list):
    rows = []
    for var in vars_list:
        sub = data[[group_col, var]].copy()
        sub[var] = pd.to_numeric(sub[var], errors="coerce")
        for g, d in sub.groupby(group_col, dropna=False):
            x = d[var].dropna()
            rows.append({"Variable": var, "Cohort": g, "N": int(x.shape[0]),
                         "Mean": float(x.mean()) if len(x) else np.nan,
                         "SD": float(x.std(ddof=1)) if len(x)>1 else np.nan})
    return pd.DataFrame(rows)

tab = summarize_by_group(df, "Cohort_Label", VARS)
table4 = (tab.pivot(index="Variable", columns="Cohort", values=["Mean","SD","N"])
            .sort_index())
# compact formatting
def fmt(m,s,n):
    if pd.isna(m): return "–"
    return f"{m:.2f} ({'–' if pd.isna(s) else f'{s:.2f}'}), n={n}"
out = pd.DataFrame(index=table4.index)
for c in table4.columns.levels[1]:
    out[str(c)] = [fmt(table4.loc[i,("Mean",c)], table4.loc[i,("SD",c)], int(table4.loc[i,("N",c)]))
                   for i in table4.index]

with pd.ExcelWriter(OUT_PATH, engine="xlsxwriter") as xlw:
    out.to_excel(xlw, sheet_name="Table4")

print("Saved:", OUT_PATH)
```

H.3 Propensity Score Matching with Caliper and ATTs

```
# Propensity Score Matching (1:1, without replacement) with caliper = 0.2 SD of LOGIT(PS)
# Files: University_Math_Scores_4_covariates.xlsx
import pandas as pd, numpy as np
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import NearestNeighbors
from scipy import stats
import statsmodels.api as sm

INPUT_PATH = "University_Math_Scores_4_covariates.xlsx"
ID_COL = "Student ID"
TREAT_COL = "Augmented"
COVARIATE_COLS = ["Math Grd12 Fin Rslt","Eng Grd12 Fin Rslt","NBT QL Score","NBT Math Score"]
OUTCOME_COLS = ["Test 1S","Test 2S","Exam S","Final S","Test 1F","Test 2F","Exam F","Final F"]

# 1) Propensity scores from the 4 covariates
df = pd.read_excel(INPUT_PATH)
df.columns = [c.strip() for c in df.columns]

X = df[COVARIATE_COLS].copy()
for c in X.columns:
    X[c] = pd.to_numeric(X[c], errors="coerce").astype(float).fillna(X[c].median())
y = df[TREAT_COL].astype(int).values

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X.values)
logit = LogisticRegression(max_iter=1000, solver="lbfgs").fit(X_scaled, y)
ps = np.clip(logit.predict_proba(X_scaled)[:,1], 1e-6, 1-1e-6)
df["pscore"] = ps
df["logit_ps"] = np.log(ps/(1-ps))

# 2) Caliper on logit(PS)
caliper = 0.2 * df["logit_ps"].std(ddof=1)

# 3) Greedy 1:1 matching treated->control within caliper (no replacement)
treated_idx = df.index[df[TREAT_COL]==1].to_list()
control_idx = df.index[df[TREAT_COL]==0].to_list()

nn = NearestNeighbors(n_neighbors=1).fit(df.loc[control_idx, ["logit_ps"]].values)
used_controls = set()
pairs = []
for i in treated_idx:
    dist, j_idx = nn.kneighbors(df.loc[[i], ["logit_ps"]].values, n_neighbors=1, return_distance=True)
    j = control_idx[j_idx[0][0]]
    if dist[0][0] <= caliper and j not in used_controls:
        pairs.append((i, j, float(dist[0][0])))
```

```
    if dist[0][0] <= caliper and j not in used_controls:
        pairs.append((i, j, float(dist[0][0])))
        used_controls.add(j)

# 4) Build matched sample
rows = []
for k,(i,j,d) in enumerate(pairs, start=1):
    t = df.loc[i].copy(); c = df.loc[j].copy()
    t["pair_id"], t["role"], t["match_distance"] = k, "treated", d
    c["pair_id"], c["role"], c["match_distance"] = k, "control", d
    rows += [t, c]
matched = pd.DataFrame(rows).reset_index(drop=True)

# 5) Balance: ASMD before/after (on the 4 covariates)
def asmd(x1, x0):
    s2 = 0.5*(np.var(x1, ddof=1) + np.var(x0, ddof=1))
    return (np.mean(x1) - np.mean(x0)) / np.sqrt(s2 + 1e-12)

def balance_table(full_df, matched_df, covs):
    out = []
    T = full_df[TREAT_COL].astype(int).values
    for c in covs:
        xb, x0b = full_df.loc[T==1, c].astype(float), full_df.loc[T==0, c].astype(float)
        xm_t = matched_df.loc[matched_df["role"]=="treated", c].astype(float)
        xm_c = matched_df.loc[matched_df["role"]=="control", c].astype(float)
        out.append({"covariate": c,
                    "ASMD_before": asmd(xb, x0b),
                    "ASMD_after":  asmd(xm_t, xm_c)})
    tab = pd.DataFrame(out)
    tab["|ASMD_before|"] = tab["ASMD_before"].abs()
    tab["|ASMD_after|"]  = tab["ASMD_after"].abs()
    return tab

bal = balance_table(df, matched, COVARIATE_COLS)
```

```
# 6) ATT for outcomes (treated - control within each pair)
def att_pairs(mdf, outcomes):
    rows = []
    for ycol in outcomes:
        diffs = []
        for pid, g in mdf.groupby("pair_id"):
            gt = g[g["role"]=="treated"][ycol].astype(float)
            gc = g[g["role"]=="control"][ycol].astype(float)
            if gt.size==1 and gc.size==1 and np.isfinite(gt.iloc[0]) and np.isfinite(gc.iloc[0]):
                diffs.append(gt.iloc[0] - gc.iloc[0])
        diffs = np.array(diffs, dtype=float)
        n = diffs.size
        if n==0:
            rows.append({"outcome": ycol, "N_pairs": 0, "ATT": np.nan, "SE": np.nan, "t": np.nan, "p": np.nan})
            continue
        att = diffs.mean()
        se = diffs.std(ddof=1)/np.sqrt(n) if n>1 else np.nan
        t  = att/se if (se is not np.nan and se>0) else np.nan
        p  = 2*(1-stats.t.cdf(abs(t), df=n-1)) if (t is not np.nan and n>1) else np.nan
        rows.append({"outcome": ycol, "N_pairs": int(n), "ATT": att, "SE": se, "t": t, "p": p})
    return pd.DataFrame(rows)

att = att_pairs(matched, OUTCOME_COLS)

# --- print a quick summary ---
print({
    "treated_total": int((df[TREAT_COL]==1).sum()),
    "control_total": int((df[TREAT_COL]==0).sum()),
    "matched_pairs": int(len(pairs)),
    "caliper_logit": float(caliper),
})
print("\nBalance (|ASMD| before→after):")
print(bal[["covariate","|ASMD_before|","|ASMD_after|"]])
print("\nATT results (first few):")
print(att.head())
```

H.4 Sensitivity Analysis

```
import pandas as pd, numpy as np, os
from scipy import stats

MATCHED_PATH = "outputs/psm_caliper_0p2.xlsx"
OUT_PATH = "outputs/sensitivity_rosenbaum.xlsx"
os.makedirs("outputs", exist_ok=True)

OUTCOME_COLS = ["Test 1S","Test 2S","Exam S","Final S","Test 1F","Test 2F","Exam F","Final F"]

xls = pd.ExcelFile(MATCHED_PATH)
matched = pd.read_excel(xls, "matched_sample")

def rosenbaum_bounds_wilcoxon(diffs, gammas):
    diffs = np.asarray(diffs, dtype=float)
    diffs = diffs[np.isfinite(diffs)]
    diffs = diffs[diffs != 0.0]  # drop zeros (no rank)
    if diffs.size == 0:
        return pd.DataFrame({"Gamma": gammas, "p_upper": [np.nan]*len(gammas),
                             "W":[np.nan]*len(gammas),"R":[np.nan]*len(gammas),"n_pairs":[0]*len(gammas)})
    signed = np.sign(diffs)
    ranks = stats.rankdata(np.abs(diffs), method="average")
    W = float(np.sum(ranks[signed > 0]))
    R = float(np.sum(ranks))
    out = []
    for g in gammas:
        pi = g / (1.0 + g)            # worst-case
        mu = pi * R
        var = pi * (1.0 - pi) * np.sum(ranks**2)
        z = (W - mu) / np.sqrt(var + 1e-12)
        p_upper = 1.0 - stats.norm.cdf(z)  # upper-tail bound
        out.append({"Gamma": float(g), "p_upper": float(p_upper), "W": W, "R": R, "n_pairs": int(diffs.size)})
    return pd.DataFrame(out)

coarse = np.round(np.concatenate([[1.0], np.arange(1.05, 3.01, 0.05)]), 2)
fine   = np.round(np.concatenate([[1.0], np.arange(1.01, 2.01, 0.01)]), 2)

summary_rows = []
with pd.ExcelWriter(OUT_PATH, engine="xlsxwriter") as xlw:
    for ycol in OUTCOME_COLS:
        diffs = []
        for pid, g in matched.groupby("pair_id"):
            gt = g[g["role"]=="treated"][ycol].astype(float)
            gc = g[g["role"]=="control"][ycol].astype(float)
            if gt.size==1 and gc.size==1 and np.isfinite(gt.iloc[0]) and np.isfinite(gc.iloc[0]):
```

```
            if gt.size==1 and gc.size==1 and np.isfinite(gt.iloc[0]) and np.isfinite(gc.iloc[0]):
                diffs.append(gt.iloc[0] - gc.iloc[0])
        diffs = np.array(diffs, dtype=float)

        if diffs.size == 0:
            pd.DataFrame({"Gamma": coarse, "p_upper": np.nan}).to_excel(xlw, sheet_name=f"{ycol}_curve", index=False
            summary_rows.append({"outcome": ycol, "n_pairs": 0, "critical_Gamma": np.nan})
            continue

        curve_coarse = rosenbaum_bounds_wilcoxon(diffs, coarse)
        crit = np.nan
        mask = curve_coarse["p_upper"] > 0.05
        if mask.any():
            crit = float(curve_coarse.loc[mask, "Gamma"].iloc[0])
            fine_grid = fine[fine <= min(2.0, crit + 0.1)]
            curve_fine = rosenbaum_bounds_wilcoxon(diffs, fine_grid)
            mask2 = curve_fine["p_upper"] > 0.05
            if mask2.any():
                crit = float(curve_fine.loc[mask2, "Gamma"].iloc[0])
            curve = pd.concat([curve_coarse, curve_fine]).drop_duplicates("Gamma").sort_values("Gamma")
        else:
            curve = curve_coarse.copy()

        curve.to_excel(xlw, sheet_name=f"{ycol}_curve", index=False)
        summary_rows.append({"outcome": ycol, "n_pairs": int((diffs!=0).sum()), "critical_Gamma": crit})

    pd.DataFrame(summary_rows).to_excel(xlw, sheet_name="summary", index=False)

print("Saved:", OUT_PATH)
```

H.5 Variation Across Subgroups

```
# Overlap-weighted WLS by NSC Math bands (70–84 vs 85–100)
# - Computes overlap weights (w = 1-p for treated, w = p for controls)
# - Fits WLS (HC1) with interaction: Y ~ treat + band_high + covars + treat:band_high
# - Outputs: tau & SE per band, p, Holm-adjusted p (per band), and p_interaction

import pandas as pd
import numpy as np
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler
import statsmodels.api as sm
from statsmodels.stats.multitest import multipletests

# --- paths & columns ---
INPUT_PATH = "University_Math_Scores_4_covariates.xlsx"   # adjust if needed
ID_COL = "Student ID"
TREAT_COL = "Augmented"
NSC_COL = "Math Grd12 Fin Rslt"

COVARIATE_COLS = [
    "Math Grd12 Fin Rslt",
    "Eng Grd12 Fin Rslt",
    "NBT QL Score",
    "NBT Math Score",
]

OUTCOME_COLS = [
    "Test 1S","Test 2S","Exam S","Final S",
    "Test 1F","Test 2F","Exam F","Final F",
]

# --- load & define bands ---
df = pd.read_excel(INPUT_PATH, sheet_name=0)
df.columns = [c.strip() for c in df.columns]
df[NSC_COL] = pd.to_numeric(df[NSC_COL], errors="coerce").astype(float)

def nsc_band2(x):
    if 70 <= x <= 84:  return "70-84"
    if 85 <= x <= 100: return "85-100"
    return np.nan

df["nsc_band2"] = df[NSC_COL].apply(lambda v: nsc_band2(v) if pd.notna(v) else np.nan)
dsub = df[df["nsc_band2"].notna()].copy()  # keep only the two bands
```

```
# --- build band indicator for interaction (1 = 85–100) ---
dsub["band_high"] = (dsub["nsc_band2"] == "85-100").astype(int)

def run_outcome(outcome):
    """
    WLS (HC1): Y ~ treat + band_high + covars + treat:band_high
    Effect in 70–84  = coef[treat]
    Effect in 85–100 = coef[treat] + coef[treat:band]
    """
    # design matrix
    X = pd.concat([
        dsub[[TREAT_COL, "band_high"]].astype(float),
        dsub[COVARIATE_COLS].apply(pd.to_numeric, errors="coerce").astype(float)
    ], axis=1)
    X["treat_x_band"] = X[TREAT_COL] * X["band_high"]
    X = sm.add_constant(X)

    y = pd.to_numeric(dsub[outcome], errors="coerce").astype(float)
    w = dsub["w_ow"].astype(float)

    mask = y.notna() & np.isfinite(y)
    X2, y2, w2 = X.loc[mask], y.loc[mask], w.loc[mask]

    # require variation in treatment and band
    if X2[TREAT_COL].nunique() < 2 or X2["band_high"].nunique() < 2:
        return None

    fit = sm.WLS(y2, X2, weights=w2).fit(cov_type="HC1")

    # effect in 70–84 (reference band)
    tau_low = fit.params.get(TREAT_COL, np.nan)
    se_low  = fit.bse.get(TREAT_COL, np.nan)
    p_low   = fit.pvalues.get(TREAT_COL, np.nan)

    # effect in 85–100 = coef[treat] + coef[treat_x_band]
    L = np.zeros((1, len(fit.params)))
    idx = list(fit.params.index)
    L[0, idx.index(TREAT_COL)] = 1.0
    if "treat_x_band" in idx:
        L[0, idx.index("treat_x_band")] = 1.0
    ttest = fit.t_test(L)
    tau_high = float(ttest.effect[0])
    se_high  = float(ttest.sd[0])
    p_high   = float(ttest.pvalue)

    p_inter = fit.pvalues.get("treat_x_band", np.nan)
```

```
    # Ns per band (non-missing Y)
    n_low  = int(((dsub["nsc_band2"]=="70-84")  & mask).sum())
    n_high = int(((dsub["nsc_band2"]=="85-100") & mask).sum())

    return {
        "outcome": outcome,
        "N_total": int(mask.sum()),
        "N_70_84": n_low,
        "N_85_100": n_high,
        "tau_70_84": tau_low,   "SE_70_84": se_low,   "p_70_84": p_low,
        "tau_85_100": tau_high, "SE_85_100": se_high, "p_85_100": p_high,
        "p_interaction": p_inter,
    }

# --- run across outcomes ---
rows = []
for ycol in OUTCOME_COLS:
    r = run_outcome(ycol)
    if r is not None:
        rows.append(r)

res = pd.DataFrame(rows).sort_values("outcome").reset_index(drop=True)

# --- Holm corrections (separately per band across outcomes) ---
if not res.empty:
    ok_low  = res["p_70_84"].notna()
    ok_high = res["p_85_100"].notna()
    if ok_low.any():
        _, p_holm_low, _, _  = multipletests(res.loc[ok_low, "p_70_84"].values, method="holm")
        res.loc[ok_low,  "p_holm_70_84"]  = p_holm_low
    else:
        res["p_holm_70_84"] = np.nan
    if ok_high.any():
        _, p_holm_high, _, _ = multipletests(res.loc[ok_high, "p_85_100"].values, method="holm")
        res.loc[ok_high, "p_holm_85_100"] = p_holm_high
    else:
        res["p_holm_85_100"] = np.nan

# --- display & (optionally) save ---
pd.set_option("display.float_format", lambda v: f"{v:.6g}")
print(res)
```

H.6 Performance on Higher-Order Questions

```
import pandas as pd, numpy as np, re
from scipy import stats
from statsmodels.stats.multitest import multipletests

MATCHED_PATH = "outputs/psm_caliper_0p2.xlsx"  # from your PSM step
HOQ_FILES = {
    "Test_1S": "Test_1S_scores.xlsx",
    "Test_2S": "Test_2S_scores.xlsx",
    "Exam_S":  "Exam_S_scores.xlsx",
    "Test_1F": "Test_1F_scores.xlsx",
    "Test_2F": "Test_2F_scores .xlsx",  # note the space
    "Exam_F":  "Exam_F_scores.xlsx",
}
ID_COL = "Student ID"

matched = pd.read_excel(MATCHED_PATH, sheet_name="matched_sample")

def parse_total_points(cols):
    for c in cols:
        for pat in [r"Total\s*\(([\d\.]+)\s*pts?\)", r"Total\s*\(([\d\.]+)\s*point[s]?\)"]:
            m = re.search(pat, str(c), re.IGNORECASE)
            if m: return c, float(m.group(1))
    raise ValueError("No 'Total (... pts/points)' column found.")

def hoq_pct_from_file(path, label):
    df = pd.read_excel(path, sheet_name=0)
    df.columns = [c.strip() for c in df.columns]
    tot_col, denom = parse_total_points(df.columns)
    pct_col = f"{label}_HOQ_pct"
    df[pct_col] = (100.0 * pd.to_numeric(df[tot_col], errors="coerce") / denom).clip(0, 100)
    return df[[ID_COL, pct_col]]

# Merge HOQ %
from functools import reduce
frames = [hoq_pct_from_file(p, lbl) for lbl,p in HOQ_FILES.items()]
hoq = reduce(lambda a,b: a.merge(b, on=ID_COL, how="outer"), frames)

m = matched.merge(hoq, on=ID_COL, how="left")
hoq_cols = [c for c in m.columns if c.endswith("_HOQ_pct")]

def att_from_pairs(mdf, outcome_col):
    diffs = []
    for pid, g in mdf.groupby("pair_id"):
        gt = g[g["role"]=="treated"][outcome_col].astype(float)
        gc = g[g["role"]=="control"][outcome_col].astype(float)
        if gt.size==1 and gc.size==1 and np.isfinite(gt.iloc[0]) and np.isfinite(gc.iloc[0]):
```

```
def att_from_pairs(mdf, outcome_col):
    diffs = []
    for pid, g in mdf.groupby("pair_id"):
        gt = g[g["role"]=="treated"][outcome_col].astype(float)
        gc = g[g["role"]=="control"][outcome_col].astype(float)
        if gt.size==1 and gc.size==1 and np.isfinite(gt.iloc[0]) and np.isfinite(gc.iloc[0]):
            diffs.append(gt.iloc[0] - gc.iloc[0])
    diffs = np.array(diffs, dtype=float)
    n = diffs.size
    if n == 0:
        return {"outcome": outcome_col, "N_pairs": 0, "ATT": np.nan, "SE": np.nan, "t": np.nan, "p": np.nan}
    att = diffs.mean()
    se  = diffs.std(ddof=1)/np.sqrt(n) if n>1 else np.nan
    t   = att/se if (se is not np.nan and se>0) else np.nan
    p   = 2*(1 - stats.t.cdf(abs(t), df=n-1)) if (t is not np.nan and n>1) else np.nan
    return {"outcome": outcome_col, "N_pairs": int(n), "ATT": float(att), "SE": float(se), "t": float(t), "p": float

att = pd.DataFrame([att_from_pairs(m, c) for c in hoq_cols]).sort_values("outcome")
if att["p"].notna().any():
    ok = att["p"].notna()
    _, p_holm, _, _ = multipletests(att.loc[ok, "p"].values, alpha=0.05, method="holm")
    att.loc[ok, "p_holm"] = p_holm
```

H.7 Higher-Order Questions (Variation Across Sub-Groups)

```python
# Overlap-weighted WLS on HOQ outcomes by NSC Math bands: 70–85 vs >85–100
# Assign 85 to the lower band (to avoid overlap): low = [70,85], high = (85,100]
import pandas as pd, numpy as np, re
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler
import statsmodels.api as sm
from statsmodels.stats.multitest import multipletests

# --- paths ---
COV_PATH = "University_Math_Scores_4_covariates.xlsx"
HOQ_FILES = {
    "Test_1S": "Test_1S_scores.xlsx",
    "Test_2S": "Test_2S_scores.xlsx",
    "Exam_S":  "Exam_S_scores.xlsx",
    "Test_1F": "Test_1F_scores.xlsx",
    "Test_2F": "Test_2F_scores .xlsx",  # note the space
    "Exam_F":  "Exam_F_scores.xlsx",
}

ID_COL = "Student ID"
TREAT_COL = "Augmented"
NSC_COL = "Math Grd12 Fin Rslt"
COVARIATE_COLS = ["Math Grd12 Fin Rslt","Eng Grd12 Fin Rslt","NBT QL Score","NBT Math Score"]

# --- helpers for HOQ % ---
def parse_total_points(cols):
    for c in cols:
        for pat in [r"Total\s*\(([\d\.]+)\s*pts?\)", r"Total\s*\(([\d\.]+)\s*point[s]?\)"]:
            m = re.search(pat, str(c), re.IGNORECASE)
            if m: return c, float(m.group(1))
    raise ValueError("No 'Total (... pts/points)' column found.")

def hoq_pct_from_file(path, label):
    df = pd.read_excel(path, sheet_name=0)
    df.columns = [c.strip() for c in df.columns]
    tot_col, denom = parse_total_points(df.columns)
    pct_col = f"{label}_HOQ_pct"
    df[pct_col] = (100.0 * pd.to_numeric(df[tot_col], errors="coerce") / denom).clip(0, 100)
    return df[[ID_COL, pct_col]]

# --- load base & banding (70–85 vs >85–100) ---
base = pd.read_excel(COV_PATH)
base.columns = [c.strip() for c in base.columns]
base[NSC_COL] = pd.to_numeric(base[NSC_COL], errors="coerce").astype(float)
```

```
def band_70to85_vs_gt85(x):
    if pd.isna(x): return np.nan
    if 70 <= x <= 85: return "70-85"
    if x > 85 and x <= 100: return "85-100"
    return np.nan

base["nsc_band2"] = base[NSC_COL].apply(band_70to85_vs_gt85)
data = base[base["nsc_band2"].notna()].copy()

# --- merge HOQ % ---
from functools import reduce
frames = [hoq_pct_from_file(p, lbl) for lbl,p in HOQ_FILES.items()]
hoq = reduce(lambda a,b: a.merge(b, on=ID_COL, how="outer"), frames)
data = data.merge(hoq, on=ID_COL, how="left")
HOQ_COLS = [c for c in data.columns if c.endswith("_HOQ_pct")]

# --- propensity scores and overlap weights ---
X = data[COVARIATE_COLS].copy()
for c in X.columns:
    X[c] = pd.to_numeric(X[c], errors="coerce").astype(float).fillna(X[c].median())
y = data[TREAT_COL].astype(int).values

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X.values)
ps = LogisticRegression(max_iter=1000, solver="lbfgs").fit(X_scaled, y).predict_proba(X_scaled)[:,1]
ps = np.clip(ps, 1e-6, 1-1e-6)
data["w_ow"] = np.where(data[TREAT_COL]==1, 1-ps, ps)
data["band_high"] = (data["nsc_band2"] == "85-100").astype(int)

def run_outcome(outcome):
    X = pd.concat([data[[TREAT_COL, "band_high"]].astype(float),
                   data[COVARIATE_COLS].apply(pd.to_numeric, errors="coerce").astype(float)], axis=1)
    X["treat_x_band"] = X[TREAT_COL] * X["band_high"]
    X = sm.add_constant(X)

    y = pd.to_numeric(data[outcome], errors="coerce").astype(float)
    w = data["w_ow"].astype(float)
    m = y.notna() & np.isfinite(y)
    X2, y2, w2 = X.loc[m], y.loc[m], w.loc[m]
    if X2[TREAT_COL].nunique() < 2 or X2["band_high"].nunique() < 2:
        return None
    fit = sm.WLS(y2, X2, weights=w2).fit(cov_type="HC1")

    tau_low = fit.params.get(TREAT_COL, np.nan)
    se_low  = fit.bse.get(TREAT_COL, np.nan)
```

```
    p_low   = fit.pvalues.get(TREAT_COL, np.nan)

    L = np.zeros((1, len(fit.params)))
    idx = list(fit.params.index)
    L[0, idx.index(TREAT_COL)] = 1.0
    if "treat_x_band" in idx:
        L[0, idx.index("treat_x_band")] = 1.0
    ttest = fit.t_test(L)
    tau_high = float(ttest.effect[0]); se_high = float(ttest.sd[0]); p_high = float(ttest.pvalue)

    p_inter = fit.pvalues.get("treat_x_band", np.nan)

    n_low  = int(((data["nsc_band2"]=="70-85") & m).sum())
    n_high = int(((data["nsc_band2"]=="85-100") & m).sum())

    return {"outcome": outcome, "N_total": int(m.sum()), "N_70_85": n_low, "N_85_100": n_high,
            "tau_70_85": tau_low, "SE_70_85": se_low, "p_70_85": p_low,
            "tau_85_100": tau_high, "SE_85_100": se_high, "p_85_100": p_high,
            "p_interaction": p_inter}

rows = []
for ycol in HOQ_COLS:
    r = run_outcome(ycol)
    if r is not None: rows.append(r)
res = pd.DataFrame(rows).sort_values("outcome").reset_index(drop=True)

# Holm-adjust per band across HOQ
if not res.empty:
    ok_low  = res["p_70_85"].notna()
    ok_high = res["p_85_100"].notna()
    if ok_low.any():
        _, p_low, _, _  = multipletests(res.loc[ok_low, "p_70_85"], method="holm")
        res.loc[ok_low, "p_holm_70_85"] = p_low
    if ok_high.any():
        _, p_high, _, _ = multipletests(res.loc[ok_high, "p_85_100"], method="holm")
        res.loc[ok_high, "p_holm_85_100"] = p_high
```