ASSESSING THE IMPROVEMENT IN LOGICAL REASONING OF STUDENTS ENROLLED IN “NUMBERS FOR LIFE” COURSE AT MCMASTER UNIVERSITY

By

MICHAELA KELLY, B.Sc.H

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

MASTER OF SCIENCE (2023)	McMaster University
(Mathematics and Statistics)	Hamilton, Ontario

TITLE: Assessing Improvement in Logical Reasoning of Students Enrolled in "Numbers for Life" Course at McMaster University
AUTHOR: Michaela Kelly, B.Sc.H
SUPERVISOR: Dr. Miroslav Lovric
NUMBER OF PAGES: 60

Abstract

To be numerate is to have the ability to understand numbers and be confident with numeric information presented in day-to-day situations. The way numeracy is defined varies between researchers; however, most agree that having skills in numeracy is essential to function in the world. In order to provide students with the opportunity for exposure to basic numeracy skills, McMaster University's course Math 2UU3 – "Numbers for Life" is offered to non-mathematics major students in second year or above. To measure the effectiveness of this course, and to determine whether students retain the numeracy skills and knowledge acquired in the course, we developed a series of assessments with questions based on content learned throughout the semester. Students were tested three times – once before completing the course, once after completing the course, and once again a year later. This study focuses in on the logical reasoning aspect of numeracy which includes understanding logical structures and being able to work through problems rationally and systematically. The results from the study reveal that students who took the course and participated in completing the given assessments showed improvement with their logical reasoning skills significantly.

Acknowledgements

Thank you to the Department of Mathematics and Statistics at McMaster for helping me succeed in mathematics – a subject that is as important as it is equally beautiful, and for funding my work over the last two years.

I would like to thank my friends – the ones I made during my master's, and the ones who have been there long before I started. They know who they are, and each of them make the heavy things in life feel lighter. I am immensely thankful to Julie Jenkins whose guidance and support over the last two years helped make this thesis possible.

Thank you to Dr. Erin Clements and Dr. Caroline Junkins for serving on my Master of Science thesis defense committee. I could not be luckier to have them both present.

Nobody uplifts me more than my family. My two sisters who I am fortunate to call my best friends inspire me every day. My parents – mom and dad, have supported me my whole life and will undoubtedly continue to do so. I thought about one person in particular constantly throughout this degree – my mom, who although could not be here physically to watch me finish my Master's, was able to begin it with me which I am especially grateful for. I dedicate this thesis to her.

Finally, a heartfelt thank you to Dr. Miroslav Lovric for being a wonderful teacher, mentor, supervisor, and friend. From the moment I started my undergraduate degree at McMaster University six years ago, Dr. Lovric has been a constant source of inspiration, encouragement, and support. It is unthinkable that I would be in this position right now – successfully completing my Master's degree, without his guidance.

Table of Contents

Introduction

While becoming quantitatively literate is a skill that needs to be learned and practiced, there needs to be just as much emphasis on the ways in which these skills are being introduced to students. "Numbers for Life" course has been designed to address the lack of numeracy knowledge and related skills among post-secondary students. Very high demand for the course suggests that contrary to some beliefs, teaching numeracy in K-12 education only does not suffice.

Although hard to be conceptualized precisely, quantitative literacy (QL) has many definitions from different domains with similar ideas. The Mathematical Association of America defines QL as "the ability to adequately use elementary mathematical tools to interpret and manipulate quantitative data and ideas that arise in individuals' lives." The importance of quantitative literacy is widely agreed upon, however there is inconsistency among what people believe it really means.

Quantitative literacy was first used synonymously with numeracy in the UK where authors aimed to find a word that would serve as "a mirror image of literacy." The Cockroft report from 1982 suggests that for one to be numerate, they must satisfy two attributes. The first attribute is the ability to utilize existing mathematical skills for dealing with practical mathematical needs of everyday life. The other attribute is the ability to appreciate and understand information when it is presented in a mathematical way, such as graphs, charts, tables, or percentages. Because of the varying opinions on the meaning of quantitative literacy, educators who are tasked with teaching courses in QL are left without substantial guidance on what methods or content should be used.

Another way to define quantitative literacy has been suggested by the Association of American Colleges and Universities (2010) who state that quantitative literacy is a "'habit of mind', competency, and comfort in working with numerical data." They believe that individuals with strong abilities in quantitative literacy are better suited in solving problems in authentic day to day scenarios, as well as understanding how to use quantitative evidence to form arguments.

Numbers are like written language in the sense that they are a type of technology of communication. Just as the human brain has the capacity to learn spoken languages, each person is born with some sort of numerate ability. This can be as simple as counting on your fingers. The emergence of exposure to quantitative literacy depended on methods of comprehension, such as the invention of numbers and written language, being sustained by social factors (Fisher, 2011).

Based on a definition by Andre DiSessa, literacy is a "socially widespread patterned deployment of skills and capabilities in a context of material support to achieve valued intellectual ends." This definition can be translated as the quantitative counterpart, where we consider the utilization and comprehension of numbers with the material support of numbers, decimals, arithmetic, graphs, calculators, computers, etc. This raises the question of which skills are necessary to utilize these numbers. The list of skills necessary is constantly changing and it varies between contexts, and so DiSessa improves his definition to say literacy is a "convergence of a large number of genres and social niches on a common, underlying representational form." Researchers have concluded that in order to be quantitatively literate, one must be numerically confident, appreciate mathematics, be able to interpret data, have the ability to think and reason logically, and apply mathematics to different contexts (Karaali, 2016).

Despite the variations of the definition of numeracy, each of these definitions aim to clarify the vagueness of what "skills required to function in today's world" really means. Essential Skills Ontario (2012) states that numeracy should be divided into five components which are reflected in day-to-day situations:

1. Math involving money or the ability to make financial transactions.
2. Scheduling, budgeting, accounting, or planning for the best use of money and time.
3. Measuring quantities, areas, volumes, and distances.
4. Analyzing data.
5. Numerical estimation.

Numeracy is often used synonymously with quantitative literacy, or sometimes "mathematical literacy" because they share common features which involve mathematics knowledge and skills. They aim to prepare individuals to function effectively and be comfortable using numbers in work and society and include the ability to apply mathematical knowledge to real-world situations. Mathematics and numeracy are not the same, but some mathematics (such as arithmetic, basic geometry, and probability) is imperative as a strong basis for applying mathematical skills in a practical world.

In simplest terms, numeracy can be thought about as a combination of specific knowledge and skills, which are needed to function in the modern world. It involves reasoning related to numeric information which can be presented in a variety of ways (numeric, graphic, and dynamic forms). In addition to the common questions associated within traditional mathematical contexts (such as: What is this? Why is this true? How do we know?), we expand numeracy to include critical, evidence-supported thinking, common sense, and logical reasoning in situations/contexts where, numbers or quantitative information are not explicitly present. Even though many authors do not mention it, logical reasoning and arguing are implicitly present in their understanding of numeracy. For instance, a numerate person should be able to understand this typical statement from a financial document: “unless a customer keeps their balance under the maximum spending limit and pays the minimum payment or more, they will be charged a fee according to the schedule on the next page.”

To be numerate does not mean that one must be proficient in advanced mathematical concepts. Instead, it is sufficient to be able to thoughtfully apply fundamental skills in mathematics that have been learned at early grade levels into real world tasks – such as filing taxes, having the correct dosage of a medication, or understanding how a mortgage works. In a similar way to literacy, which is constantly being developed, numeracy is also a skill that is constantly being refined. Many authors do agree that the way mathematics is being taught in schools focuses too heavily – or in some cases exclusively - on abstract concepts and processes and does not adequately cover authentic real-world problems. According to Brooks (2013), when teaching numeracy, material should be presented “in context for a specific purpose; reason for learning is to solve a problem and apply it.”

Numeracy is a very important skill to have, yet a survey done by the Conference Board of Canada in 2012-2014 revealed that 55% of Canadian adults do not have sufficient numeracy skills, and the number has grown significantly within the last decade. Although adequate numeracy skills among adults is very low, it is likely not as prevalent among university students. Despite this, it is still a serious concern. The need for numeracy skills is evident in greater society, as we are faced with important, complex, and sometimes ambiguous and vaguely stated questions and issues in which it is important to inform ourselves. These questions and issues are inclusive but not limited to elections, freedom of speech, societal inequalities, public investments, health and well-being, etc. Numeracy is also essential in recognizing and evaluating the information we see in texts and online from which we can form meaningful and evidence-supported opinions.

Research Questions

This thesis is part of a larger, government funded (HEQCO = Higher Education Quality Council of Ontario) research project which aims to determine the extent of which the course "Numbers for Life" taught at McMaster University develops numeracy skills among students to evaluate the effectiveness of the way the course is taught and to study how it contributes to the development of numeracy skills. This project fits into HEQCO's mission of evaluating efforts to improve numeracy knowledge and to develop and upgrade transferable skills among university students, and within the general population.

As the course "Numbers for Life" aims to prepare students for the numeracy demands of day-to-day life, it is important that we make sure students are engaged with the content and are genuinely benefiting from the material learned. For my thesis, I aim to answer the following question(s):

(R1) To what extend does the course, "Numbers for Life" develop students' numeracy skills? More specifically, how has students' ability to use logical reasoning (such as understanding and creating cause and effect arguments) and recognizing logical fallacies changed after completing the course, and are students able to better communicate their answers to quantitative reasoning problems and questions after completing the course?

(R2): To what extent are numeracy skills retained a year after the course is complete?

(R3): How do students' academic backgrounds, interests and experiences affect the development of their numeracy skills?

The research questions were inspired by the importance of numeracy in general, as well as by the absence of research into similar courses in other universities, because there are very few similar courses in Canada. Thus, positive outcomes of this research could be used to promote the development and teaching of numeracy courses on a broader scale in Canada and beyond.

Literature Review

The primary objective of this thesis is to measure the effectiveness of the way "Numbers for Life" is taught at McMaster University and to investigate the ways the course contributes to an overall improvement in students' learning of numeracy as well as their development of numeracy skills. In particular, we focused in on the logical aspect of these skills and how they were developed over the span of taking "Numbers for Life."

Despite there being an extensive number of research on numeracy and its benefits, assessment of numeracy, and numeracy instruction, there does not currently exist many publications which address the specifics of numeracy instruction. Further, the body of literature that measures the ways in which students need to argue logically in day-to-day scenarios is very limited. It is a common occurrence in mathematics education research that only a small number of researchers study post-secondary classroom teaching practices. In addition, mathematics education researchers are, as a rule, not involved in authoring textbooks for these courses, nor in creating online learning materials.

The main questions we are exploring - how do we teach numeracy, and how do we know that we have succeeded (or not) - are poorly covered in literature. Our research aims to fill this gap, by providing insights into the teaching of specific topics (see *Research Questions*) and assessing to what extent students have learnt these topics. Another aspect of our research is that we assess retention, by testing students one year after they completed the course. There is one publication we know of that addresses numeracy being learned in the longer term. A study by Mandell & Klein (2009) assesses the long-term effects of numeracy education by comparing two groups of students – one group which took a specific financial literacy course, and another group which did not. The results of this study showed that there was ultimately no difference in the two groups of students in terms of being financially literate. Further, students who were enrolled did not find themselves to be any better at saving money than the students who never took the course. As a result of these findings, the study raised some questions about the effectiveness of financial literacy courses in the longer term.

The importance of the development of numeracy knowledge and skills among both students and adults has been recognized in research literature, as well as in various government documents (Dingwall, 2000; Dion, 2014; Geiger, Goos, & Forgasz, 2015; National Numeracy, 2019; Orpwood & Sandford Brown, 2015). Despite this, it is still a challenge to effectively teach numeracy at all levels of formal education. Research that supports this challenge of teaching is slim and inadequate. Geiger, Goos, & Forgasz (2011) believe that there is not much known about how those who teach numeracy learn about, appropriate, and create effective mathematics teaching practices.

While some believe that numeracy education requires its own space for instruction, others claim that it should be integrated within traditional education and taught as a supplement in many subject areas. In his paper that defines quantitative literacy, Fisher (2019) concludes that it is "the facility to participate in the intersecting quantitative practices of many different communities, each with their own patterns of discourse." As literacy is innately social, courses that address numeracy need to provide students with the opportunity to engage in discussions which involve numbers. These discussions should also aim to encompass many different areas in which mathematics presents itself. The social contexts where numbers are found should be the focus, more than the list of skills required when preparing future instructors to teach courses on quantitative literacy. Due to numeracy appearing naturally in a broad number of disciplines, instructors from these disciplines (including non-mathematicians) should try to introduce numeracy practices into their curriculums.

We contribute to this debate by analyzing teaching and learning in a university-level numeracy course, which places numbers in multiple contexts and authentic real-life situations. This approach aligns with the ideas of Steen (2001); however, we diverge in the implementation: Steen believes that numeracy must be addressed in all (school, university) subjects (i.e., through an integrated interdisciplinary curriculum), and that there is no need for a specific numeracy course. Recognizing the realities of slow decision-making processes and occasional aversion toward significant curricular changes within universities, rather than wait, we created a free-standing numeracy course at McMaster University, which is open to all students on our campus (except to Math and Stats majors).

Among the research addressing teaching of numeracy, we consulted those that were specific in studying teaching practice, and creating resources, such as Goos (2016). Within the Australian context, Kissane (2012) discusses numeracy-related projects in school mathematics. Several authors suggest word problems (Gravemeijer, 1997; Karaali, 2008; Hoogland & Pepin, 2016) as the means of facilitating the development of the problem-solving skills, albeit not always in authentic, real-life contexts.

Financial literacy is explored in a study by Sunderaraman, Barker, Chapman and Cosentino (2020) in which a population of 88 cognitively healthy adults were used to examine the relevance of various aspects of numeracy for financial literacy. Results of this study show that numeracy plays a major role in making financial decisions and it is important for researchers and instructors to understand how the different aspects of numeracy relate to financial literacy. Moreover, proficient numeracy appears to be a key predictor of various financial outcomes such as investment and retirement decisions.

Durst and Kaschner (2019) explore student performance on true or false assessments involving logical implication statements of the conditional form "If P, then Q" with the intent to better understand the ways in which students process conditional statements in logic, and to see if their logical misconceptions are impeding their ability to demonstrate mathematical knowledge. The study involved the administering an online assessment to students enrolled in a Calculus II course. Findings from this study suggest that students do in fact make errors on the true or false items, but that the errors are due to a misunderstanding of logical structures and unrelated to how well they can grasp concepts in calculus.

Kaye (2014) discusses teaching environments and suggests an interesting departure from teaching numbers (within numeracy) and, instead, focussing on topics such as data and measuring shape and space.

The assessment of numeracy knowledge and skills, on its own and not as part of assessment of mathematical knowledge and skills, is relatively new. Geiger, Goos, & Forgasz (2015) offer an overview of assessment efforts, both on a large scale (international assessments) and on smaller, local scales.

Various types of mathematical reasoning in mathematics education were outlined in a systematic review done by Hjelte, Schindler, and Nilsson (2020). The review provided an overview of the different ways people reason through mathematic problems and what kinds of methods of reasoning are addressed in empirical research in mathematics education. This study is only an overview, as with limited resources, it is hard to generalize and account for every single type of mathematical reasoning. The results of this study show that scholars either have a domain-general, or a domain-specific view on mathematical reasoning. In the domain-general view, reasoning is not bound to a specific mathematical topic, and it can be applied in any mathematical task or subject. This conclusion is essential to numeracy development because in order to be numerate, one must be able to apply fundamental mathematics ideas and calculations to real-world contexts.

Tout and Gal (2015) write about the creation of assessment frameworks for two major assessments of numeracy skills: Programme for International Student Assessment (PISA) and the Programme for International Assessment of Adult Competencies (PIAAC, also known as the OECD Survey of Adult Skills). Both frameworks have been the result of collaboration of several international experts. Of course, the relationship between assessment and instruction is not linear. The authors write: "while assessments per se should not drive instruction, the PISA and PIAAC frameworks do reflect cumulative wisdom and research findings that are important to examine in a systematic manner."

In creating assessment instruments for this project, we were guided by the assessment frameworks for PISA and PIAAC, as well as published research on numeracy assessment, questions from various surveys of numeracy, and materials that aim to prepare students for numeracy tests. Of course, we narrowed our focus on questions that helped us to answer our research questions.

Gaze et al. (2014) offer a thorough walk-through the development of their quantitative literacy/quantitative reasoning (QLR) assessment instrument. Their definition of QLR as "the skill set necessary to process quantitative information and the capacity to critique, reflect upon, and apply quantitative information in making decisions" is close to generally accepted conceptualizations of numeracy. Gaze et al. (2014) offer insights into the assessment development, that we found quite valuable, and include suggestions to:

- Replace procedural, algorithmic questions with more involved reasoning, critical thinking questions.
- Ask students to interpret tables and charts rather than doing it for them.
- Focus on quantitative literacy, using numbers in meaningful sentences rather than just computation.
- Ask students to postulate possible explanations for statistics rather than traditional logic games.

To analyze the difficulty of our assessment tasks, we consulted Weingarten, Brumwell, Chatoor & Hudak (2018) numeracy proficiency levels.

There are many sources (such as MyWellesley, n.d.; Hoogland & Pepin, 2016) which offer, what they claim to be, numeracy tasks. However, often, one must go through them carefully to separate those that focus on mathematical rather than numeracy skills (such as performing arithmetic operations which do not come from a real-life context.

Methodology

In this study, our goal was to evaluate the retained numeracy knowledge and skillset of students who had been enrolled and completed the numeracy-based math course "Numbers for Life" at McMaster University. To do this, we needed to observe how well these skills were being used in both the short term and long term.

Our survey population consists of the Fall 2021 cohort of students who were enrolled in Math 2UU3, "Numbers for Life" at McMaster University. About two thirds (384/592) of the students gave consent for their data to be used in our study, which constitutes a very large sample size.

The first piece of data collected that we felt may be important in answering the research questions was information on students' academic backgrounds. We organized students by the faculty they were earning a degree in. This consisted of the following degree programs: Health Science, Science, Commerce, Engineering (students in this faculty were predominantly in computer science engineering), Humanities, Social Science, and Arts & Science. This information aided us in answering our research question (R3) which explores how student backgrounds and experiences influence their success in this particular course.

Next, we used course assessments which were collected from the Fall 2021 cohort.

The pre- and post-tests given to the consenting students contained the same questions and were administered at the start of the course (pre-test), at the end of the course (post-test), and one year later (delayed post-test). We studied students' responses to 13 numeracy-based questions which related to topics that would have been familiar to them from the course, or possibly from their previous schooling or from life experiences. As a novelty in this kind of research, we measured a longer-term retention of these skills by surveying the Fall 2021 cohort once again in December 2022. Due to the difficulty of reaching out to students who are no longer part of a course for their participation, we incentivized them by offering gift card of a small value to thank them.

In order to measure the improvement in skills and knowledge, we could not rely only on quantitative comparisons of pre and post-test scores. Instead, we looked closer at students' work to better understand how their learning and development of skills had changed. As much of the data collected was going to be analyzed qualitatively, we developed rubrics. The rubrics used were in accordance with research questions 1 and 2 and spanned the following categories:

Category A: A student's ability to understand numbers (e.g., relative size, absolute size, numerical patterns) and work with numeric information (e.g., approximate, scale, visualize, estimate).

Category B: A students' ability to use logical reasoning (such as understanding and creating cause and effect arguments) and recognize logical fallacies.

Category C: A students' ability to engage with multiple step-problems which require quantitative reasoning.

Category D: A student's ability to communicate their answers to quantitative reasoning problems and questions.

In this thesis, I focus mainly on the questions which required an understanding of the structure of logical sentences and logical reasoning, so the category that was most useful in qualitatively analyzing those responses was category B, with category D being useful in determining how well their answers were communicated.

To use the rubrics, we developed a mechanism to score how well students were able to logically reason through the problems, and if they were able to clearly communicate and justify their thought process.

This study is quite unique in that we were able to collect data three times to generate a repeated experiment. This was the only way to keep the research controlled and consistent, as new university student cohorts often differ significantly from year to year. By repeating this experiment, we are able to control the effects of pedagogical interventions that may affect the success of different cohorts, as well has distinguish between outcomes which would be consistent and those which would have only occurred in one cohort.

We also conducted delayed post-test interviews with students from the Fall 2021 cohort. Due to limited time, we interviewed only six students, and we did not analyze the responses to the delayed post-test questions which were given to all the students to respond to. The interviewed students were informed of the purpose, benefits, and risks of the study (*see Appendix D)* and were given 60 minutes to answer the questions (*see Appendix B*). A more thorough study of the data that we have will be conducted starting in Fall 2023.

We have obtained ethics clearance with McMaster University Research Ethics Board. Recruitment was accomplished through McMaster Ethics Board-approved means: class announcements, advertisements on the course web page, and individual emails to students. The letter of information and consent form were introduced in a lecture, posted on the course web page, and mailed to each student. A graduate research assistant ensured, through standard ethics-approved routines (such as coding of names), that students' privacy is fully protected.

Results

Pre- and Post-Test Evaluations

In order for us to answer our research questions on the numeracy skills developed, and how students are able to communicate their quantitative reasoning, we asked students to complete two surveys comprising of 13 numeracy-based questions (*See Appendix A*). The first survey was given near the beginning of the semester to the students who were enrolled in "Numbers for Life" in Fall 2021. In addition, a large group of those students were asked to complete the same survey at the end of the course. This allowed us to compare their responses, to determine how well they answered the questions to get a sense of their improvement. We evaluated the responses using a rubric, and in this thesis, we are looking specifically at the improvements in logical reasoning and communication in responses.

While logic presents itself in some way in all areas of mathematics, logical reasoning is not something that is centred out as a specific concept in mathematics education. Logic is in some capacity used in solving every math problem, even something as simple as adding two single digit numbers. Unlike topics related to numbers and developing number sense, logical reasoning is not taught explicitly in secondary school, or in post-secondary courses except for specific courses in mathematics and philosophy. For this reason, it is likely that any improvements in performance on these logical-based questions are very likely due to the course instruction in "Numbers for Life."

Post-test numerical data was normalized in order to provide a more meaningful comparison to the pre-test. The normalization was based on the amount of time it took students to complete the full pre-test and post-test. However, this does not resolve the issue completely, as it is impossible to know all the reasons why students finished the test quicker during the post-test in comparison to their pre-test. Reasons for this could include the fact that students could finish faster because they had learnt the material, but the lack in detail in some of the responses in the post-test that had previously been given in the pre-test dispute this. Despite not having found any literature to support this, we still felt it was an appropriate figure to represent scores and measure improvement.

The average time it took students to complete the pre-test was 137.5 minutes with the range being between 5 minutes to nearly 24 hours, and the average completion time for the post-test was 116.89 minutes with a range between 1 minute and nearly 24 hours. By calculating the ratio between these values, we found

$$\frac{137.5}{116.89} = 1.176$$

This figure tells us that students, on average, spent 17.6% more time on completing the pre-test than they did on the post-test. The normalizing factor used to mitigate the effects of this was half of this increase (i.e., 8.8%) and helped us obtain the numbers in the last column of Table 1.

Question	Rubric Category	Pre-Test (Out of 5)	Post Test (Out of 5)	Normalized Post Test
11				
	Category B: Logical Reasoning (**Logical Structures**)	3.49	4.04	4.40
	Category D: Ability to Communicate with Quantitative Information (**Communication of Solutions**)	3.45	4.54	4.94
	Category D: Ability to Communicate with Quantitative (**Explanations of Quantitative Reasonings**)	2.99	3.14	3.42
12				
	Category B: Logical Reasoning (**Logical Structures**)	3.32	3.67	3.99
	Category D: Ability to Communicate with Quantitative Information (**Communication of Solutions**)	3.25	3.17	3.45
	Category D: Ability to Communicate with Quantitative (**Explanations of Quantitative Reasonings**)	3.02	2.66	2.89

Table 1: Mean values on questions about logical reasoning, marked using rubrics on a scale from 0-5, together with the normalized pre-test means.

Questions 11 and 12, which aimed to evaluate students' understanding and the use of mathematical logic and logical reasoning, were graded using rubrics (*see Appendix C*). There is clearly a significant improvement in logical reasoning as well as ability to communicate this reasoning on question 11. This is evident by the increase in post-test scores as both a raw score and a normalized score. While the ability to logically reason through a problem improved from pre-test to post-test in question 12, we did not see the same improvement in the assessment of ability communicate quantitative information.

Question	Topic	Pre-Test	Post-Test	Delayed post-test
6				
	Distinguishing between causation and correlation in the case of symptoms and disease	225/384 (58.59%)	249/374 (66.58%)	65/119 (54.62%)
7				
	Venn diagram of a relationship between two populations	349/384 (90.89%)	337/374 (90.11%)	103/119 (86.55%)

Table 2: Comparison of the Fall 2021 cohort of students on the pre-test and post-test on multiple choice questions (note: questions 7 was changed from multiple choice to an "explain" question in the delayed post-test).

Question 6 which highlighted the understanding of causation and correlation saw a significant increase in performance, even for the sample size being slightly smaller. Question 7 did not have much of a change between pre-test and post-test, as both means were extremely high. In the delayed post-test, question 7 was adjusted so that students were required to choose the correct Venn diagram and then justify their reasoning. With time being a limiting factor for our research, we were not able to go through these responses and grade them for a thorough qualitative analysis. The intent in changing the format of the question was to help us obtain data which is more telling in revealing how well students understood numeracy concepts and how effectively they communicated their answers.

Delayed Post-Test Interviews

The only source of delayed post-test information we were able to use for this thesis were from 6 interviews conducted in early March 2023. For the purpose of this thesis focusing primarily on logical reasoning, we will only analyze and discuss the results of questions 5, 6, 7, and 8 (*see Appendix B*). The responses to these questions were verbal, and students were assessed based on their general correctness and ability to justify their reasoning when being questioned further on their thought process.

Question	Topic	Result (based on overall correctness)
5		
	Venn diagram of a relationship between two populations	5/6 (84%)
6		
	Distinguishing between causation and correlation in the case of symptoms and disease	6/6 (100%)
7		
	Ability to produce the negation of a statement	5/6 (84%)
8		
	Recognizing different conditional statements	3/6 (50%)

Table 3: Results of the delayed post-test interviews conducted in March 2023, results showing how many students out of a sample size of 6 answered correctly. Not much importance should be placed on the bare statistics as the sample size is too small.

The results from these delayed post-test interviews were very well articulated for questions related to logical reasoning, and the ability to have a conversation with the students via the verbal interview allowed us to ensure students were being thorough in how they communicated their responses. Question 7 is interesting in the sense that there are two ways to disprove the statement, and one can use either of the two ways, or both, to show that it is false. Similar to how responses were marked in the written the Fall 2021 pre-test, and the written post-test in Fall 2022, students were given a score of 4 if they mentioned only one of two possible ways to disprove the statement; a student who mentioned only one way to disprove the statement in their interview received a half mark.

Additionally, students who could correctly conclude that the conditional statements given in question 8 were not reinterpretations of each other got partial credit if they were not able to fully justify their answer.

Discussion

The results of administering the survey instruments and interviews and qualitatively analyzing responses showed us that there was an overall improvement in numeracy skills related to logical reasoning of students who took the course "Numbers for Life." This points in a positive direction in answering our first research question which aims to determine to what extent are numeracy skills and knowledge learned and retained after completing the course.

Pre-Test and Post-Test Responses

I analyzed four questions which were related primarily to testing students on their improvement in understanding logical statements, with a secondary focus on how well they were able to communicate their responses. Two questions in the pre and post-test were formatted as multiple choice responses, and so were not graded as extensively and not in as much qualitative detail as the two of the questions which allowed students to respond by providing some justification for their thought process.

In the first written response-style question (*see Question 11, Appendix A*) which assessed the improvement in logical reasoning, there was an overall increase in correct responses shown in the post-test. Students were asked how they would disprove the statement:

In my neighbourhood there are ten dogs and they all bark at night.

The statement takes on the logical form "A and B," where A= "there are ten dogs" and B= "they all (i.e., all dogs) bark at night." In order to disprove a statement like this logically, you would need to show that either component (or both components) of the statement is not true – either A is false, or B is false, or both A and B are false.

In other words, students would need to claim that to disprove the statement, there would either need to be more or less than ten dogs, or there would need to be one dog that did not bark at night. Students who were able to clearly state that disproving both components independently as well as provide a clear justification on why received full credit based on the rubric created to assess their competency (*see Appendix C*). Here is a sample of two responses:

> "First, I would check if there are 10 dogs. Then I would then have to determine if each dog is barking at night by checking to see if all 10 are barking at night. Through this if you find that the amount of dogs is more or less than 10 then the statement is false and if there are not exactly 10 dogs barking at night the statement will be false." – 2439 (pre-test) (Bachelor of Science, Life Sciences)

> "You would have to prove that there are more or less than 10 dogs in the neighborhood. Alternatively, you could also prove that not all 10 bark at night by finding one that does not bark even once in the night. For this statement to be true both clauses of the statement need to be true and disproving even one clause would render it false." – 90580 (pre-test) (Bachelor of Science, Biology)

The two responses above are clearly communicated. We see that the two students can correctly articulate their justification and show that they understand how to correctly negate a conjunction statement.

Next. We look at the following responses:

> "The neighbourhood doesn't have 10 dogs (can be more or less) OR - the neighbourhood has 10 dogs, but they don't all bark at night. If I use the structure learnt in class: in my neighbourhood there aren't 10 dogs, or they don't all bark at night." – 91017 (post-test) (Bachelor of Science, Life Sciences)

> "To disprove it: not (A and B) = not A OR not B. In my neighbourhood, there are not 10 dogs, or they don't all bark at night." – 93558 (post-test) (Bachelor of Arts, French and Linguistics)

Each of these responses demonstrates the understanding that to disprove the statement, negating either of the components in this statement will suffice. Students who argued that one would only need to disprove one part of the statement were correct but scored a 4 (out of 5) on the rubric category B (which measured understanding of logical structures) for not recognizing every possible way to disprove it. Student #91017 even mentions in their post-test answer that they used the structure that they had specifically remembered learning in class. While some students may struggle to remember the specifics of what they learned in the course, they are able to recognize familiar words or phrases that were brought up in class at some point. This, however, may suggest that students only remember certain words or topics, but do not internalize that information effectively enough to be able to apply it in the short term.

A theme that showed up in both pre-test and post-test (but more frequently in the pre-test) was the idea of disproving the statement by gathering tangible evidence. As numeracy is about authentic situations, we take such answers seriously, even if they do not involve traditional, expected, numeric approaches. About 97 among the 384 students in the pre-test answered the question by stating they would be able to show there were not 10 dogs in the neighbourhood by conducting experiments, collecting video evidence, or listening for 10 distinct barks by recording. Some examples of these evidence-based responses are given:

> "To prove that this statement is not true, it would be best to put each of the 10 dogs under observation for one night. The observational footage would then be examined the next morning to see if indeed all 10 dogs were barking the previous night or not". – 38355 (pre-test) (Bachelor of Science, Life Sciences)

> "I would have to find evidence to prove this statement wrong by doing repeated experiments. For example, I can monitor the 10 dogs every night for a span of two weeks. If I find that not dogs DO NOT all bark every night, then I can prove the statement wrong. The statement in the question is not a mathematical theorem that has been proven or established, but a real life situation, which is not clear-cut. Real life differs from pure mathematics in the sense that repeated experiments with similar outcomes can be accepted. Although math theory cannot be built on statements that have not been proven rigorously, real life statements can. Thus, if in real life, if there is evidence that NOT all 10 dogs bark every night, then this would prove the statement above as not true". – 18838 (pre-test) (Bachelor of Health Sciences)

> "We would have to monitor each individual dog either by physically being there and looking at them or through cameras to see whether each of the 10 dogs is barking at night or whether a few are barking a lot giving the impression that all 10 are simultaneously barking. If all dogs aren't barking at night, the statement will be proven to be false." – 89507 (pre-test) (Bachelor of Health Sciences)

Responses that were similar to those above were less common in the post-test, where only about 14 students relied on some visual or audible evidence-based means to successfully make the statement false. An interesting trend found in these responses is that nearly every student that gave an evidence-based solution related the evidence to the second part of the conjunction – "all 10 dogs bark at night". It was less common for students to want to gather evidence on the number of dogs in the neighbourhood – such as finding out whether there were indeed only 10 dogs in the neighbourhood.

Where some students lacked sufficient detail in communicating their thought process, it was evident that they still understood how to work with logical statements. Most responses that I concluded exhibited latent evidence were very short responses which lacked sufficient detail. (Latent evidence is referring to the meaning underlying what is said or shown.) Thus, we can infer that the following students know how to correctly disprove the statement:

> "Firstly, go to the neighbourhood to check if there are 10 dogs. Then, stay up and count how many dogs bark at night. Verify all of them do so." – 10905 (pre-test) (Bachelor of Engineering, Computer Science)

> "How does he know there are 10 dogs?" – 69697 (post-test) (Bachelor of Arts, Social Sciences)

While these answers are concise and do not provide much detail, there is some latent evidence to suggest students understand how disproving a conjunction works. Student 10905 mentions to go and check if there are 10 dogs but does not state what it would mean if there were more or less than 10 dogs. Similarly, as both components of the statement are briefly mentioned – verifying the number of dogs in the neighbourhood and checking they all bark, the student is aware that they could show either of those parts are not true to successfully disprove the statement. Student 69697 asks a follow up question to the original. By asking how we know for certain there are 10 dogs, the student may be alluding to the fact that if there was some way to prove there were not actually 10 dogs, then the statement could be disproved completely.

The overall increase in performance of responses to question 11 reveal a promising result on how "Numbers for Life" developed student's numeracy skills pertaining to logical reasoning. Responses in the post-test showed a less-common occurrence of evidence-based responses, and I found that more students were likely to either negate the quantifier of the statement or find one dog that did not bark. The following responses include both of these attempts to disprove the statement:

> “Some possible ways this statement could be proven not to be true are counting the number of dogs in the neighbourhood and finding that there are less than 10 dogs or more than 10 dogs or finding at least one dog that does not bark at all during the night.” – 20031 (post-test) (Bachelor of Science, Life Sciences)

> “Either find that there are not 10 dogs in the neighbourhood (find out how many dogs there are in the neighbourhood) OR prove that one of the 10 dogs do not bark at night.” – 24255 (post-test) (Bachelor of Science, Psychology, Neuroscience, and Behaviour)

> “I need to prove that in my neighbourhood, there are not 10 dogs (can be more than 10 dogs or can be less than 10 dogs) or not all dogs bark at night. Satisfying one of the situations (not 10 dogs / not all dogs bark at night) can disprove the statement.” – 20238 (post-test) (Bachelor of Commerce)

> “You would have to prove either the assumption or the outcome is not true. For example, if there are not exactly 10 dogs in the neighbourhood, the statement is not true by contradiction. Further, if there are 10 dogs in the neighbourhood you will need to prove at least one does not bark at night.” – 54840 (post-test) (Bachelor of Arts, Social Sciences)

While I viewed a complete answer as the four responses shown above, I was also glad to see that many students in the post-test were able to at least mention one of the ways to disprove the statement – by either negating the quantifier, or by only finding a single dog that does not bark. The following responses illustrate this:

> “To show that there exists at least one dog that doesn't bark at night.” – 8378 (post-test) (Bachelor of Arts, Social Sciences)

> “Show that there are more or less than 10 dogs.” – 38571 (post-test) (Bachelor of Health Sciences)

> "You will simply have to find one of those 10 dogs and prove that it is not barking. Finding one dog that is not barking in the neighborhood would mean the statement is not true." – 89507 (post-test) (Bachelor of Health Sciences)

> "To prove the statement is not true, we should find that there is one of these 10 dogs doesn't bark at night." – 36168 (post-test) (Bachelor of Commerce)

Disproving only one of the parts of a conjunction is enough to disprove the entire statement, so these students are almost entirely correct. It would be interesting, in a future study, to question students when testing them on statements like this by asking them if they would disprove it the same way if the statement were a disjunction instead of a conjunction. Some information pertaining to this notion was gathered during the delayed post-test interviews, which will be discussed later in this thesis.

The second question (*see Question 12, Appendix A*) that required a written solution and probed students' understanding of logical structures involved conditional statements. When comparing the pre-test performance to the post-test performance, there was a slight increase in the logical reasoning category B. The same positive results were not observed in category D that measured communication. This question took the following statements:

If you do not study, you will not pass the test. **(1)**

If you do study, you will pass the test. **(2)**

Students were asked if the two above statements were correct reinterpretations of each other, and then to explain why or why not.

The statements are related to each other as negative statements ("negative" meaning that the constituent parts of the two statements are negatives/ negations of each other). The first statement (1) is of the form "If A, then B" where A = "you do not study," and B = "you will not pass the test." Statement (2) takes on the form of the original conditional statement, "If not A, then not B" where not A = "you do study," and not B = "you will pass the test."

This question also encourages students to demonstrate their understanding of causation versus correlation, in particular, knowing that correlation does not imply causation. The following responses give some insight into how students understand this notion in the pre-test:

> "Saying both of these statements assumes that there is one that causes the other. While this seems logical, it isn't always true that if one studies, they will pass the test. There is however, a strong positive correlation between studying and passing tests." – 91017 (pre-test) (Bachelor of Science, Life Sciences)

> "No, this is not a correct conclusion. The first sentence represents a cause and effect. Studying is the cause and passing the test is the effect. This is a causality. If the causality is true, then the converse may or may NOT be true. In other words, just because not studying will cause someone to fail a test, does not mean that studying will cause them to pass the test for sure. Of course, studying increases the chances of passing a test. However, studying varies per student, and inefficient studying or studying for very little amounts can also lead to failure. Thus, studying can still lead to failing a test, so the conclusion in the question is incorrect." – 18838 (pre-test) (Bachelor of Health Sciences)

> "No this is not a correct conclusion because studying and passing is correlation not causation. Studying for a test does not cause you to pass the test. They are positively correlated meaning if you study you are more likely to pass, however it is not guaranteed that if you study you will pass." – 29633 (pre-test) (Bachelor of Science, Kinesiology)

> "No, this statement is not correct. The original statement declares that not studying causes failing the test. Though, it does not establish a relationship between studying and passing the test. It is reasonable to assume that there is a correlation between studying and passing, but that does not necessarily mean that in every instance of study will the student pass the test. Therefore, this reinterpretation is not accurate". – 57039 (pre-test) (Bachelor of Health Science)

These responses, among many others, show that students already had some idea that causation does not imply correlation before taking the course. With the success rate of logical reasoning skills improving by the scores from the rubric, it is likely that many students who completed the course learned about these ideas for the first time. This is because logical reasoning is not explicitly taught in high school math courses, it is instead used implicitly within math problems. This may especially be the case for many students enrolled in "Numbers for Life." Since the course is not open to mathematics majors, these students have probably not taken more than first year calculus, if they have taken that at all.

There was a common occurrence of students often confusing what negative and converse statements meant. In both the pre-test and post-test, many students seemed to be unaware that the converse of a statement of the form "If A, then B" is "If B, then A". A few examples of this are given:

> "This is an incorrect conclusion because if a causality is true, the converse may or may not be true. People who don't study for tests typically fail however not everyone that studies for a test will pass their test." – 48555 (pre-test) (Bachelor of Commerce)

> "No this is not true. The converse is not always true for a saying. There may be a correlation between studying and passing or failing a test. But the fact that one studies does not cause one to pass- it is simply related." – 11677 (pre-test) (Bachelor of Commerce)

> "No, this is not a correct conclusion because "If you study, you will pass the test" is the converse and does not necessarily imply that it will be correct. There is no relation between the truthfulness of a statement and its converse. Therefore, these statements are NOT equivalent." – 442 (pre-test) (Bachelor of Science, Life Sciences)

> No, this is a converse, which is not necessarily true. However, if the statement was "if you pass the test, then you studied" then that could be an interpretation of the original statement because it is it's contrapositive." – 90738 (pos-test) (Bachelor of Commerce)

None of the above responses are wrong, because each response reveals that the student can recognize that the statements are not correct reinterpretations of each other. However, students seemingly cannot easily distinguish the difference between the converse, negative, and contrapositive of a conditional statement despite knowing when each of the different types are correct or incorrect as reinterpretations. For example, student 90738 in the post-test mistook the negative for the converse (like many other students) and was correct in stating that the converse of a true statement is not always necessarily true. They also rewrote the statement correctly as the contrapositive but did not give an answer that provided sufficient explanation to the particular question.

In a similar way that students relied on something like tangible evidence to disprove the statement in question 11, many students responded to this question with an argument that involved external factors. In the pre-test, the results comprised of many responses that argued the second statement was not a correct reinterpretation, because reasons beyond our control could contribute to passing or not passing a test. While answers of this style showed up again in the post-test, they were not as common. Some examples of responses based on external factors are shown below:

> “This is not the correct conclusion because there are many factors that go into passing the test. For example, one could study too much for the test and not sleep at all for the upcoming test. This would result in a very fatigued student, thus leading to an increased chance of failing the test.” – 26806 (pre-test) (Bachelor of Commerce)

> “No, this is not a correct conclusion because those students who didn't study can be lucky enough to guess the answers or cheat and pass the test also some may not need to study because they know the content. On the other hand, some students who did study may not pass because they might not have a grasp on the given content and simply do not understand how to achieve the correct answers.” – 28161 (pre-test) (Bachelor of Arts, Social Sciences)

> “It's not a correct conclusion because whether you can pass the exam is not just determined by whether you study hard. Although whether you study or not accounts for a large part of the proportion, at the same time, your physical condition on the day of the test and the difficulty of the test questions ext. will also affect whether you pass the test.” – 88279 (pre-test) (Bachelor of Commerce)

> “Not a correct conclusion because if you don't study, you will simply not have enough information and practice necessary to pass the test (assuming you did not have any prior knowledge of the subject). However, if you do study, it is not a guaranteed event that you will pass your test. There may be other variables that cause you to fail the test, even if you prepared well enough by studying.” – 1281 (post-test) (Bachelor of Science, Psychology, Neuroscience, and Behaviour)

> "No, this conclusion is not correct, as even if you study, there is still a chance you fail the test, for example, the test could be overly hard, or you may have studied the wrong things." – 8627 (post-test) (Bachelor of Science, Life Sciences)

This is a small sample of the many responses that argued that the reason the statements were not equivalent was because of reasons beyond our control. These responses are very human – they do not rely on logical structures and knowledge of conditional statements to justify their arguments. This may be an indication that students answer differently when they feel like they can relate to a question. Failing or almost failing a test because you did not put enough time towards studying, or because an emergency beyond your control occurred is a very common experience for a lot of students. Doing better or worse than one thought for the amount of effort one puts into a test is also common.

There is a notion of learning inside the classroom versus outside the classroom that is discussed by the Education Summary ("Difference Between Learning Inside and Outside the Classroom," 2022). Learning inside the classroom means that the focus is placed on giving knowledge and discipline indoors so that students can be more focused on their learning process. Learning outside the classroom allows students to learn through experiences and practices, where they can be challenged in a more practical setting. Students may be answering question 12 by mentioning these external factors because they find it relatable and are attributing their responses to behaviours, values, and attitudes learned outside the classroom. As mentioned earlier – although such responses might not be acceptable as an answer to a mathematical question, they are viewed as acceptable, and even desirable, in a numeracy setting.

The two multiple choice questions related to logical understanding did not require students to justify their reasoning, so there is not much to discuss in detail for either of them. Question 6 (*see Appendix A*) asked students to distinguish between correlation and causation in the case of disease and symptoms. The question reads:

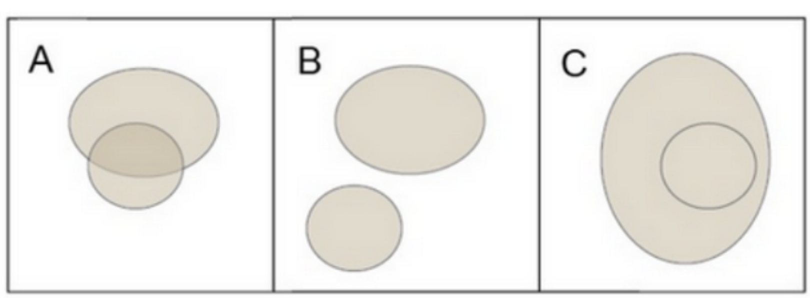
The sentence "The symptoms of meningitis are high fever, neck pain, and seizures" expresses a:

- **(a) Correlation between meningitis, and high fever, neck pain, and seizures.**
- **(b) Correlation between high fever, neck pain, and seizures.**
- **(c) Causation, with meningitis being the cause.**
- **(d) Causation, with high fever, neck pain, and seizures being the cause.**

The correct answer to this question is (c), with a fair number of students (58.59%) selecting that correct response. The second most common answer selected in both the pre-test and post-test was (a). Students may have chosen the correlation between meningitis, high fever, neck pain, and seizures since each of these events are likely to occur simultaneously, but most students were able to recognize that in order to exhibit these symptoms, there needs to be a reason or a cause for them to occur. The thought process of students answering this question is briefly explored in the interview results analysis.

The last question to be discussed in this thesis (*see question 7, Appendix A*) from the test that aims to measure logical understanding involves selecting the correct diagram to represent a statement related in inclusion and exclusion.

"Some animals with thick fur are mammals, and some mammals have thick fur." Which diagram represents the relationship between animals with thick fur and mammals?



There is not much to discuss related to this question, as the students' performance was highly successful in all tests. It seems that students were aware that there should be some common portion in the diagram that represent mammals with thick fur, along with separate portions that represent the animals with thick fur, and the mammals respectively.

Interviews with Students

The interviews conducted in March 2023 were the only piece of data that could give us an insight into answering the third research question – how well numeracy skills and knowledge have been retained in the long term. Since the sample size is so small, I could not extract much meaningful, generalizable data from them. In the very least, it indicated an overall positive set of results for the questions related to logic.

In the delayed post-test, question 5 was the Venn diagram question (*see Appendix B*). Due to the nature of the interviews, I was able to gather information on the students' thought process and how they communicated their choice. Only one student among the 6 who were interviewed selected the incorrect option to represent the given conditions. They chose option B and argued that both isolated circles represented animals with thick fur which are mammals, and mammals with thick fur respectively. The student explained that option A is incorrect because if you assume the middle section represents thick fur, the larger circle represents animals, and the smaller circle represents mammals, then the animals with thick fur that are mammals and mammals without thick fur are not accounted for. They go on to say that option C doesn't work either because if the smaller circle represents mammals with thick fur, and the outside circle is animals with thick fur, then it would suggest that animals with thick fur cannot be mammals which is a direct contradiction of the statement in the question. This student seemed frustrated by this question and stated that they found these questions to be somewhat difficult, even after the "Numbers for Life" course. Based on how they allocated the different parts of the statements to circles in the diagrams, it would suggest that the student has a fundamental difficulty in understanding inclusion and exclusion in the context of logic. This is partially confirmed with the student explaining that they think these questions may be difficult because it is hard to identify the relationship between statements. Another student mentioned their disdain for these questions as well, believing that they are easy to overthink, and that it is hard to decide which circle represents what – despite answering the question correctly.

The other 5 students chose option A, and each of them had similar reasoning for their selection, as well as reasoning for why it was not options B or C. When asked how they arrived at their conclusion, one student said:

> "There are mammals in both – so animals that have thick fur are mammals and mammals with thick fur, so I know that some of them are going to have the same attributes so it wouldn't be something like B because it has no overlap. Option C means that there could be some mammals that don't have thick fur, or some animals with thick fur that aren't mammals."

This student can effectively identify the different categories to be considered in the diagram and how to arrange them. One student considered the white area of the box surrounding the circular portions of the Venn diagram – viewing the whole diagram as opposed to the circles alone. Another student said:

> "Not all mammals have thick fur so you can't pick C because then the big category would be animals with thick fur, and then mammals would fall under it because some of them don't have thick fur. It also can't be B because this is saying that there is no correlation between animals with thick fur and mammals because some mammals do have thick fur. The only right answer would be A where one of the categories would be animals with thick fur, and the other category would be mammals, and the middle region would be mammals that have thick fur."

From the delayed post-test interviews, I gained some insight on the multiple choice question 6 (*see Appendix A*):

> **The sentence "The symptoms of meningitis are high fever, neck pain, and seizures" expresses a:**
>
> **(a) Correlation between meningitis, and high fever, neck pain, and seizures.**
>
> **(b) Correlation between high fever, neck pain, and seizures.**
>
> **(c) Causation, with meningitis being the cause.**
>
> **(d) Causation, with high fever, neck pain, and seizures being the cause.**

Among all the questions on logical reasoning in the interview, this question was the most well done with each student interviewed answering correctly. Students' justification for choosing the correct response is also consistent among participants. Students were made to think about the term "cause and effect" when they saw the word causation, and they connected the cause to meningitis and the effects to the symptoms that result from it. One student said:

> "I see meningitis, which is a condition, has symptoms of high fever, neck pain, and seizures. If we were looking at a correlation, that would just say that one thing is correlated to the other. But there is no temporal component where one thing comes after the other because you do not have high fever, neck pain, seizures, and meningitis independently. It's the meningitis that comes first and it causes these symptoms, so it's more of a causation than a correlation because the condition comes first and causes these symptoms."

Some students also mentioned that it was the phrasing of the statement that helped them decide on option (c). The key word is "of." When they read "the symptoms *of* meningitis are…", they understood it as the set of symptoms being something that belongs to meningitis or occurs because one has meningitis.

Question 7 in this delayed post-test was the question on a logical conjunction. Students were again asked how they would disprove the following statement:

In my neighbourhood there are ten dogs and they all bark at night.

Four out of the six students interviewed mentioned that to disprove the statement, you would either need to show that there were not ten dogs, or that they do not all bark at night. They explained that only showing one of those would disprove the entire statement. One student placed emphasis on what they called a key term being "all." If the statement is claiming that all ten dogs bark at night, then finding at least one dog among those ten that does not bark contradicts the part of the statement that says *all* dogs bark at night.

In order to make sure they truly understood why that was the case, I asked some students if their answer would change if the statement was changed to disjunctive form:

In my neighbourhood there are ten dogs or they all bark at night.

Students who were asked to disprove this alternate form correctly answered in saying that in this case, a statement with two conditions separated by "or" requires you to disprove both parts in order to disprove the entire statement. This gave us a positive result on retention on logical numeracy skills (but, again, they are relatively insignificant due to the small sample size).

Lastly, we questioned students to gain insight on their understanding of causation and correlation in the case of the given statements, and whether or not they were equivalent:

If you do not study, you will not pass the test. (1)

If you do study, you will pass the test. (2)

A common response to why these are not equivalent was that it is possible that you could study and still fail, and if you do study, there is a chance you can still pass the test. Students believe that you cannot take the converse or negative of a statement and claim that it is true.

As a follow up, I asked each student how they would rewrite the statement correctly in order to determine if their reasoning for thinking the statements were not equivalent was valid. Some of the responses for correctly restating "if you do not study, then you will not pass the test" were:

a. "If you study, then you may or may not pass the test."
b. "You will not pass the test if you do not study."
c. "If you do study, they you may or may not pass the test."
d. "If you do not study, then you may or may not pass the test."

Statement (a) and statement (c) are both the same answer, as partial negatives of the original statement, with ambiguous conclusions – either passing the test or not passing the test. None of the students could correctly identify the type of conditional statement this was. They either did not know or remember, or they mistook the negative of a conditional statement for the converse.

One student explicitly stated that they remembered how to solve this problem from class one year prior, and provided the reasoning that "if A, then B" where A means you do not study, and B means you do not pass the test, then the reinterpretation of the statement is the case were you have the negative of A – you do study for the test, and the conclusion of that is that you may or may not pass the test. The student attempted to give a clear explanation by breaking the statement into components but did not answer with complete correctness.

It is evident in the delayed post-test interviews, as it was in the pre and post-tests, that students were reading this question and answering based on feeling or personal experience rather than logic. I had asked students if they knew what contrapositive and negative or converse statements were independently, and they were able to correctly define their structure, however they were unable to correctly apply those structures to rephrasing the question using the contrapositive.

As part of this thesis that aims to measure the retention of numeracy knowledge and skills, we analyzed the responses that students provided when asked "What will you remember five years from now?" Most of the students (about 85-90%) stated that financial literacy stuck with them the most, and that they will remember the lessons on mortgage, credit card interest, T4, taxes, and tax brackets, understanding debt and interest, consumer price index, and inflation.

Unfortunately, not many students mentioned logical reasoning in their responses. The students who did mention it emphasized the value on using logic in real life, such as using cause and effect statements, and knowing that causality does not always hold in the converse. However, other students stated that while they may use logic in future math courses, they probably won't use it in their day to day life. This could be the result of students not seeing the value of logical understanding beyond the academic setting, or that the statements being used to instill logical understanding are not realistic enough.

> "One thing that I will remember from this course is the use of logical quantifiers as a fast way to write note or something else in our daily life. Some symbols can just replace the

> words. For example, conjunction and disjunction." – 14143 (Bachelor of Arts, Social Sciences)

> "One thing that I think I will remember from this course five years from now is how to logically work through problems. There is always a step by step approach that will help with solving the problem/issue in math or in real life scenarios." – 24147 (Bachelor of Integrated Science)

> "I will remember the unit based on logic, contrapositives, converses, etc. I feel like this logical method to thinking is very applicable to different scenarios in life and can be used in making critical and informed decisions, no matter the subject. Overall, all of the topics discussed in this course were very memorable and will most likely continue to remain relevant five years from now when I can continue to call back on them for everyday applications to my daily life and thinking." – 42020 (Bachelor of Science, Biology)

Another common answer to the question of what student will remember in 5 years is the content relating to reading small print, and the importance of understanding the different ways quantitative information is given and framed to us. However, this is not directly related to the retention of logical structures and logical reasoning.

Looking at the longer-term retention, we found a study by Mandell & Klein (2009), which showed that students who took a course on financial literacy and management were no more financially literate or had better financial behaviour by the end of those course than those who did not take the class. However, Sunderaraman, Barker, Chapman, & Cosentino (2020) found that numeracy is a key indicator of many financial outcomes including retirement and investment decisions. Unfortunately, we could not find any studies that would address the retention of the elements of logical reasoning.

Another finding was that students who did not do as well on the pre-test performed significantly better on the post-test. This revealed that students with a poorer background learned a lot during the course. We did not look closely into the backgrounds of students who took "Numbers for Life," but based on a collection of student demographic data, I found that most students who took the course came from a science or business background, and that students seeking a Bachelor of Arts degree in both Humanities and Social Sciences were less common.

Some students showed significant improvement between the pre-test and the post-test. One student with a commerce background had initially answered the question on how to disprove the statement about dogs barking (*see Question 11, Appendix A*) using evidence based methods, but in the post-test, was able to apply what they had learned about logical statements to explain how they would disprove the given statement.

> "I would go and ask the neighbors who owned the dogs and check if their dogs bark at night. If the dog does not have an owner, then I will find out myself." – 9685 (pre-test) (Bachelor of Commerce)

> "There are less or greater than 10 dogs and they are bark at night, or there are 10 dogs, but they are not all bark at night, or there are less or greater than 10 dogs and they are not all bark at night. If one of the above condition holds, the above statements is not true." – 9685 (post-test) (Bachelor of Commerce)

Not only did the student indicate the two ways to disprove the statement, but they also demonstrated their understanding through effective communication.

It was also evident that students who have a science background and are completing a Bachelor of Science degree experienced improvement after completing the course. The following responses show the difference in logical understanding of a response given by a student in the Health Sciences program:

> "I would send a person to each house during the hours of sundown to sunrise to see if at least one of them have barked every 5 mins. I believe it is very unlikely this would occur; therefore, this statement would be incorrect." – 62414 (Bachelor of Health Sciences)

> "Find one dog that doesn't bark out of the 10." – 62414 (Bachelor of Health Sciences)

While the post-test response is very concise, and lacks sufficient explanation and communication of logical reasoning, there is still the improvement of the student leaning away from a more evidence-based, impractical method to disprove. In the post-test, the student can now recognize how simple disproving the statement is when reading it in a logical manner. This tells us that students – even those in a science degree-seeking program, gained something valuable from the course.

> "This is a correct conclusion because both statements represent a positive correlation between studying and passing the test. In the first statement, the less you study, the less your chances are of passing the test. Thus, both variables are decreasing. In second statement, the more you study, the more your odds are of passing the test. Thus, both variables are increasing. For that reason, both statements can be used synonymously as they both indicate a positive correlation." – 12601 (pre-test)

> "This is not a correct conclusion because if the first statement is not true (i.e.: if you do not study, but pass the test), then the second statement may or may not be true (i.e.: if you do study, you may or may not pass the test). Since there are no clear causality between the two statements, the effects may or may not be correlated." – 12601 (post-test)

This student – who once thought the statements were equivalent, can now correctly identify that the second statement is not a correct conclusion. In the first, pre-test response, they believed the statements were equivalent and justified it using correlation. In the post-test, the student still used correlation, but was able to recognize that the statements were not equal because causation does not equal correlation – a notion covered thoroughly in Math 2UU3.

Conclusion

We began this study with the intention to determine how effective instruction in "Numbers for Life" is, and how to improve it so that similar courses can be taught within other institutions. To achieve this, we assessed students before and after completing the course to see if the improvement was present, identified by the correctness and quality of students' responses. Responses to the assessments created to evaluate students on their success of learning numeracy skills have shown us that students are genuinely retaining knowledge from the course. The main purpose of this component of the research project was to answer the first research question:

"To what extend does the course, "Numbers for Life" develop students' numeracy skills? More specifically, how has students' ability to use logical reasoning (such as understanding and creating cause and effect arguments) and recognizing logical fallacies changed after completing the course, and are students able to better communicate their answers to quantitative reasoning problems and questions after completing the course?"

Based on our findings, it is evident that students gained a better grasp on thinking and reasoning logically, but that they somewhat lacked sufficient skills to explain and communicate their reasoning. It seems that many students are comfortable with the different forms of logic in the questions they were given, but that they were, in a sense, answering the questions from a personal bias and not thinking about the answer mathematically, which was especially evident in the question about not passing a test (*see Question 12, Appendix A*).

While we could not extensively study the effects of the course long term and answer our second research question "to what extent are numeracy skills retained a year after the course is complete?", the responses to the interviews from the six chosen students showed that students could answer numeracy based questions confidently one year later and explain their reasoning well even when prompted with numerous follow up questions. These results however are not significant because the sample size is too small to provide meaningful evidence. Therefore, the delayed post-test results need to be analyzed further to draw any more evidence-based conclusions to answer this research question.

Our results showed that student background did not impede the learning of anyone who took the class, and that quality responses, as well as improvements in numeracy skills, were observed among students in various programs. This helped in answering the third research question "how do students' academic backgrounds, interests, and experiences affect the development of their numeracy skills?" Students who are enrolled in science-based programs who likely encountered calculus during their undergraduate career did not have an advantage over other students in different degree programs who did not need to take calculus in first year. Therefore, the numeracy skills taught in "Numbers for Life" are accessible to all student types and it is possible to succeed without much background in higher level mathematics.

In conclusion, the "Numbers for Life" course at McMaster University works, and it does improve numeracy skills and knowledge among students. As a result of these findings, more universities should be encouraged to adopt similar courses that teach students basic numeracy skills that will help them function in a world that is dependent on having these skills.

Bibliography

1. Borromeo Ferri, R. (2015) Mathematical Thinking Styles in School and Across Cultures. In: Cho S. (Eds.) Selected Regular Lectures from the 12th International Congress on Mathematical Education. Springer, Cham. doi: https://doi.org/10.1007/978-3-319-17187-6_9.

2. Brooks, C. (2013). Approaches to teaching adult numeracy. In Griffiths, G. & Stone, R. (Eds). *Teaching Adult Numeracy: Principles and Practice* (pp. 141-156). Open University Press, England.

3. The Conference Board of Canada (n.d.; the page claims that the data is accurate as of June 2014). Adults With Inadequate Numeracy Skills. http://www.conferenceboard.ca/(X(1)S(lqndmbt2qinu3shxmz12tk1c))/hcp/provincial/education/adlt-lownum.aspx?AspxAutoDetectCookieSupport=1

4. Cuoco, A., Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An organizing principle for a mathematics curriculum. Journal of Mathematical Behavior, 14(4), 375–402.

5. Difference Between Learning Inside and Outside the Classroom. Education Summary. (2022). https://educationsummary.com/lesson/utilizing-learners-experiences/

6. Dingwall, J. (2000). Improving Numeracy in Canada. National Literacy Secretariat. Ottawa: Human Resources and Social Development Canada. http://publications.gc.ca/collections/collection_2008/hrsdc-rhdsc/HS38-13-2000E.pdf.

7. Dion, N. (2014). Emphasizing Numeracy as an Essential Skill. Toronto: Higher Education Quality Council of Ontario. https://www.heqco.ca/SiteCollectionDocuments/Numeracy%20ENG.pdf.

8. Durst, S. & Kaschner, S.R. (2019). Logical Misconceptions with Conditional Statements in Early Undergraduate Mathematics. PRIMUS, vol. 30, No. 4, pp. 67-378. https://doi.org/10.1080/10511970.2019.1597000.

9. Essential Skills Ontario (2012). *Literacy and Essential Skills in Ontario*. Toronto: Author. http://www.essentialskillsontario.ca/sites/www.essentialskillsontario.ca/files/Literacy%20and%20Essential%20Skills%20in%20Ontario_Final2_0.pdf.

10. Frankel, J. R., & Wallen, N. E. (2009). How to Design and Evaluate Research in Education. New York: McGraw Hill.

11. Fisher, F. (2019). What Do We Mean By Quantitative Literacy? Shifting Context, Stable Core: Advancing Quantitative Literacy in Higher Education. The Mathematical Association of America. pp. 3-15.

12. Gaze, E. C., Montgomery, A., Kilic-Bahi, S., Leoni, D., Misener, L. & Taylor, C. (2014). Towards Developing a Quantitative Literacy/Reasoning Assessment Instrument, Numeracy: Vol. 7 : Iss. 2 , Article DOI: http://dx.doi.org/10.5038/1936-4660.7.2.4 Available at: http://scholarcommons.usf.edu/numeracy/vol7/iss2/art4.

13. Geiger, V., Goos, M. & Dole, S. (2011). Teacher Professional Learning in Numeracy: Trajectories Through a Model for Numeracy in the 21St Century. In J. Clark, et al. (Eds.), Mathematics: Traditions and [New] Practices. Proceedings of the 34th annual conference of the Mathematics Education Research Group of Australasia).and the Australian Association of Mathematics Teachers). Adelaide: AAMT and MERGA.

14. Geiger, V., Goos, M. & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: a survey of the state of the field. ZDM. 47. 531-548. 10.1007/s11858-015-0708-1.

15. Goos, M. (2016). Transforming professional practice: Teaching numeracy across the curriculum (lecture slides). The University of Queensland.

16. Gravemeijer, Koeno (1997). Solving word problems: A case of modelling? Learning and Instruction, Vol. 1. No. 4, pp. 389-397.

17. Hoogland, K. & Pepin, B. (2016). The intricacies of assessing numeracy: Investigating alternatives to word problems, ALM International Journal, Volume 11(2), pp. 14-26.

18. Karaali, Gizem (2008). Word Problems: Reflections on Embedding Quantitative Literacy in a Calculus Course. Numeracy 1, Iss. 2 (2008): Article 6. DOI: http://dx.doi.org/10.5038/1936-4660.1.2.6.

19. Karaali, Gizem, Edwin H., Villafane Hernandez, and Jeremy A. Taylor. (2016). “What’s in a Name? A Critical Review of Definitions of Quantitative Literacy, Numeracy, and Quantitative Reasoning.” Numeracy, 9(1): Article 2.

20. Kaye, D. (2014). Critical issues in numeracy practice – contradictions and strategies. Adults Learning Mathematics: An International Journal, 9(2), 54-62.

21. Kissane, B. (2012). Numeracy: connecting mathematics. In B. Kaur & T. L. Toh (Eds.), Reasoning, communication and connections in mathematics: yearbook 2012 (pp. 261–287). Singapore: World Scientific Publishing Co.

22. Manaster, A. B. (2001). Mathematics and Numeracy: Mutual Reinforcement. In (L. A. Steen, Ed.) Mathematics and Democracy; The Case for Quantitative Literacy (pp. 67-72). The Woodrow Wilson National Foundation Fellowship.

23. Mandell, L. & Klein, L. S. (2009). The Impact of Financial Literacy Education on Subsequent Financial Behavior. Journal of Financial Counseling and Planning Volume 20, Issue 1 2009.

24. My Wellesley (n.d.). Study Packet for the Quantitative Reasoning Assessment, Wellesley College. Available at: https://www.wellesley.edu/qr/qr-assessment-study-packet.

25. National Numeracy. (2019). Why is Numeracy Important? Available at: https://www.nationalnumeracy.org.uk/why-numeracy-important.

26. OECD/Statistics Canada (2005), *Learning a Living: First Results of the Adult Literacy and Life Skills Survey*, OECD Publishing, Paris, https://doi.org/10.1787/9789264010390-en.

27. Orpwood, G., & Sandford Brown, E. (2015). Closing the Numeracy Gap. http://www.numeracygap.ca/assets/img/Closing_the_numeracy_gap_V4.pdf.

28. Peters, E., Vastfjall, D., Slovic, P., Mertz, C.K., Mazzocco, K., Dickert, S. (2006). Numeracy and Decision Making. https://journals.sagepub.com/doi/10.1111/j.1467-9280.2006.01720.x.

29. Statistics Canada. (2013). Skills in Canada: First Results from the Programme for the International Assessment of Adult Competencies (PIAAC) https://www.statcan.gc.ca/pub/89-555-x/89-555-x2013001-eng.htm.

30. Steen, L. (2001). The case for quantitative literacy. In L. Steen (Ed.), Mathematics and democracy: the case for quantitative literacy (pp. 1–22). Princeton: National Council on Education and the Disciplines.

31. Tout, D., & Gal, I. (2015). Perspectives on numeracy: Reflections from international assessments. ZDM Mathematics Education, 47, 691–706. DOI 10.1007/s11858-015-0672-9.

32. Weingarten, H. P., Brumwell, S., Chatoor, K. & Hudak, L. (2018). Measuring Essential Skills of Postsecondary Students: Final Report of the Essential Adult Skills Initiative. Toronto: Higher Education Quality Council of Ontario. https://www.heqco.ca/SiteCollectionDocuments/FIXED_English_Formatted_EASI%20Final%20Report%282%29.pdf.

Appendix A
Pre-test post-test and delayed post-test questions

1. Pre-test Fall 2021, Post-test Fall 2021

A sweater costs S dollars. As the price tag offers a 13% discount, you decide to buy it. At the counter, the 13% sales tax is added to the discounted price. How much will you pay for the sweater: S dollars, less than S dollars, or more than S dollars? Explain how you arrived at your answer to this question.

Delayed Post-test December 2022

You are thinking of buying a sweater, and you look at the price tag. Next to it, a sign says that there is a 13% discount on all sweaters, and you decide to buy it. At the counter, the 13% sales tax is added to the discounted price. Will you pay more, less, or the same as the price seen on the price tag? Explain how you arrived at your answer.

2. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test December 2022

The table below states the nutritional facts for a container of milk.

Nutrition Facts	
Serving size 1 cup (220 g)	
Servings per container 3.5	
Amount per serving	
Calories 140	Calories from Fat 60
	% Daily Value
Total Fat 8 g	12%
Saturated fat 2g	9%
Trans fat 0 g	
Cholesterol 10 mg	4%
Sodium 235 mg	10%
Total Carbohydrate	10%
Dietary Fibre 0 g	
Sugars 12 g	
Protein 20 g	36%

If you drink the entire container, how many calories would you consume?

3. Pre-test Fall 2021, Post-test Fall 2021

The energy demand that a person needs for its cardiovascular system to function normally is computed by multiplying their cholesterol level by 9.5 and then dividing by their heart rate. Person A and Person B have the same heart rate, but person A has lower cholesterol level. Select the correct statement.

- ◯ Person A's energy demand is larger than Person B's energy demand
- ◯ Person A's energy demand is smaller than Person B's energy demand
- ◯ Person A's energy demand is equal to Person B's energy demand

Delayed Post-test December 2022

The energy demand a person needs for its cardiovascular system to function normally is computed by multiplying their cholesterol level by 9.5 and then dividing by their heart rate. Person A and Person B have the same heart rate, but person A has lower cholesterol level. What is the relationship between Person A's and Person B's energy demands? Show your reasoning.

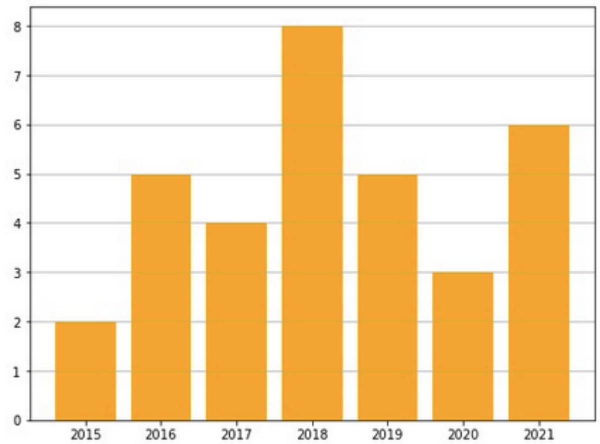
4. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

The chances of four events occurring, during a winter storm day in Ontario, are given below. Which event is the *most likely* to occur?

- ◯ 1 in 45,000 of injury in a car crash
- ◯ 10 in 400,000 of injury from slipping and falling
- ◯ 2 in 50,000 of mild to medium hypothermia
- ◯ 10 in 500,000 of injury from an exploding fireplace

5. Pre-test Fall 2021, Post-test Fall 2021

The diagram below shows the net profit (in $ thousand) of a small family-owned company.

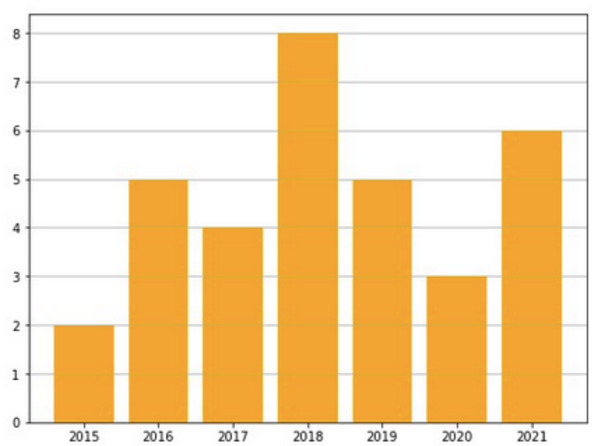


When did the company experience the largest relative growth?

- ◯ from 2015 to 2016
- ◯ from 2017 to 2018
- ◯ from 2020 to 2021

Delayed Post-test 2022

The diagram below shows the net profit (in $ thousand) of a small family-owned company.



True/False: The company experienced the largest relative growth from 2017 to 2018. State your answer as true or false, and explain your reasoning.

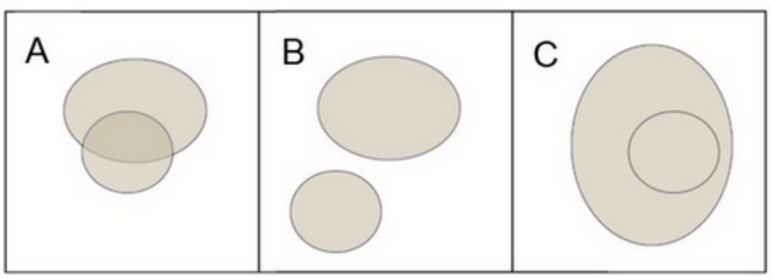
6. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

The sentence "The symptoms of meningitis are high fever, neck pain, and seizures" expresses a

- ◯ correlation between meningitis, and high fever, neck pain and seizures
- ◯ correlation between high fever, neck pain, and seizures
- ◯ causation, with meningitis being the cause
- ◯ causation, with high fever, neck pain, and seizures being the cause

7. Pre-test Fall 2021, Post-test Fall 2021

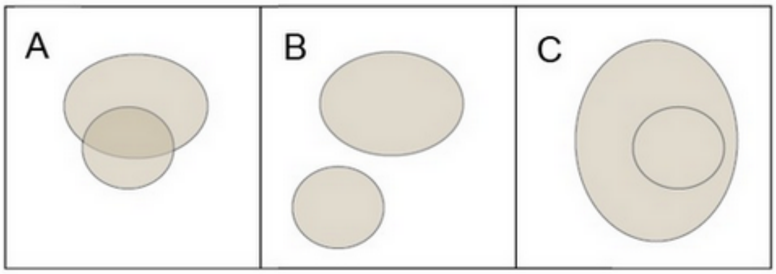
"Some animals with thick fur are mammals, and some mammals have thick fur." Which diagram represents the relationship between animals with thick fur and mammals?



- ◯ diagram A
- ◯ diagram B
- ◯ diagram C

Delayed Post-test 2022

"Some animals with thick fur are mammals, and some mammals have thick fur." Which diagram represents the relationship between animals with thick fur and mammals?



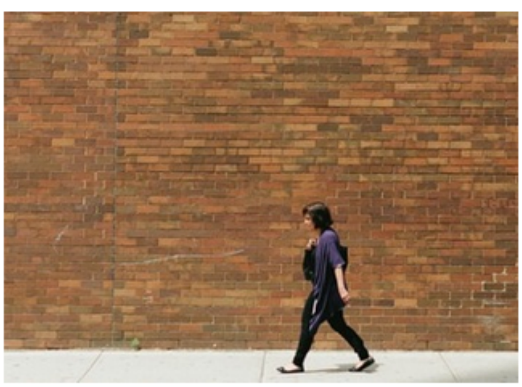
Answer by stating one of A, B, or C, and justify your reasoning.

8. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

You are in New Zealand, using your Canadian phone. Assume that a roaming charge for the usage over your plan is $6 per 80MB (megabytes) of data, and assume that you have reached the limit of your plan, meaning that you have to pay for extra data you use. You decided to watch a movie, whose size is 1.6 GB (gigabytes). How much will you pay in roaming charges for watching this movie?

9. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

Estimate the area of the part of the wall shown in this picture. (Not the area of the picture, but the area of the actual, real, wall shown below). Explain your reasoning.



10. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

Last year, there were 100 monkeys on island X and 1000 monkeys on island Y. This year, there are 200 monkeys on island X and 1100 monkeys on island Y. Thus, on both islands, the populations of monkeys increased by 100. How would you describe what is different about the change of two populations?

11. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

Consider the statement "In my neighbourhood there are 10 dogs and they all bark at night." What would you have to do to prove that this statement is not true?

12. Pre-test Fall 2021, Post-test Fall 2021, Delayed Post-test 2022

The statement "If you do not study, you will not pass the test" can be reinterpreted as "If you study, you will pass the test." Is this a correct conclusion? Why or why not?

13. Pre-test Fall 2021, Post-test Fall 2021,

Postal codes in Cook Island have the format LL-dddd where L is an uppercase letter and d is a digit; for instance, DX-3402. The first letter must be one of A, B, C, D, E, F, or G, and there are no restrictions on the second letter. The first digit in dddd cannot be 0 and cannot be 9, and there are no restrictions on the remaining three digits. What is the maximum number of postal codes available? Show how you arrived at your answer.

Delayed Post-test 2022

This ad for Harvey's burgers claims that there are *8 million ways to top your burger*.

You are suspicious of this fact, and decide to investigate. You discover that to customize your burger, you have to pick one of the three options for a bun (white, multigrain, or no bun), and then you can choose as many toppings as you wish among: 3 premium toppings, 10 other toppings and 10 sauces; for instance, you can have 2 premium toppings, 8 other toppings and 7 sauces. Based on this information you calculate the number of ways to top your burger. Does your estimate match Harvey's estimate of 8 million? Explain your reasoning.

Appendix B
Interview questions

Question 1

You are thinking of buying a sweater, and you look at the price tag. Next to it, a sign says that there is a 13% discount on all sweaters, and you decide to buy it. At the counter, the 13% sales tax is added to the discounted price. Will you pay more, less, or the same as the price seen on the price tag? Explain how you arrived at your answer.

Question 2

The energy demand a person needs for its cardiovascular system to function normally is computed by multiplying their cholesterol level by 9.5 and then dividing by their heart rate. Person A and Person B have the same heart rate, but person A has lower cholesterol level. What is the relationship between Person A's and Person B's energy demands? Show your reasoning.

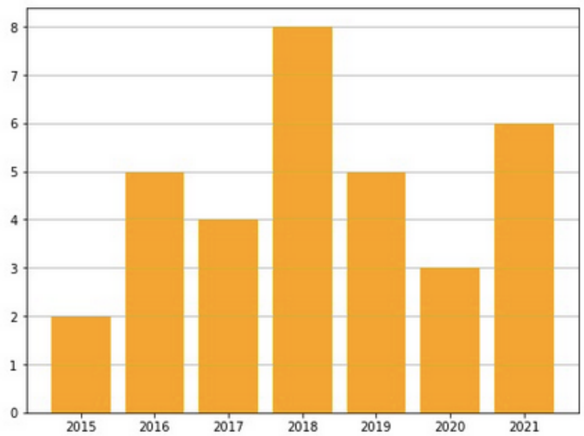
Question 3

The chances of four events occurring, during a winter storm day in Ontario, are given below. Which event is the *most likely* to occur?

- ◯ 1 in 45,000 of injury in a car crash
- ◯ 10 in 400,000 of injury from slipping and falling
- ◯ 2 in 50,000 of mild to medium hypothermia
- ◯ 10 in 500,000 of injury from an exploding fireplace

Question 4

The diagram below shows the net profit (in $ thousand) of a small family-owned company.



True/False: The company experienced the largest relative growth from 2017 to 2018. State your answer as true or false, and explain your reasoning.

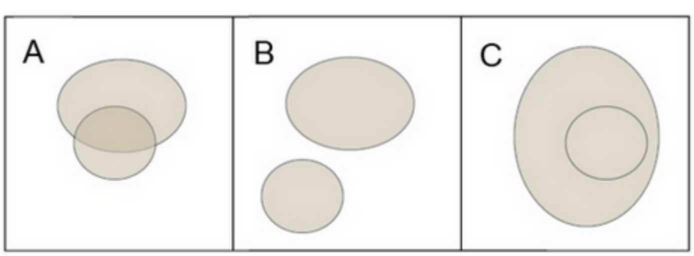
Question 5

The sentence "The symptoms of meningitis are high fever, neck pain, and seizures" expresses a

- ◯ correlation between meningitis, and high fever, neck pain and seizures
- ◯ correlation between high fever, neck pain, and seizures
- ◯ causation, with meningitis being the cause
- ◯ causation, with high fever, neck pain, and seizures being the cause

Question 6

"Some animals with thick fur are mammals, and some mammals have thick fur." Which diagram represents the relationship between animals with thick fur and mammals?



Answer by stating one of A, B, or C, and justify your reasoning.

Question 7

Consider the statement "In my neighbourhood there are 10 dogs and they all bark at night." What would you have to do to prove that this statement is not true?

Question 8

The statement "If you do not study, you will not pass the test" can be reinterpreted as "If you study, you will pass the test." Is this a correct conclusion? Why or why not?

Question 9

This ad for Harvey's burgers claims that there are *8 million ways to top your burger*.

You are suspicious of this fact, and decide to investigate. You discover that to customize your burger, you have to pick one of the three options for a bun (white, multigrain, or no bun), and then you can choose as many toppings as you wish among: 3 premium toppings, 10 other toppings and 10 sauces; for instance, you can have 2 premium toppings, 8 other toppings and 7 sauces. Based on this information you calculate the number of ways to top your burger. Does your estimate match Harvey's estimate of 8 million? Explain your reasoning.

Appendix C
RUBRICS FOR GRADING PRE/POST-TEST RESPONSES

The objective of this phase of research was to develop an appropriate assessment structure for explanatory and qualitative student responses to the pre-, post-, delayed post- and in-course assessments administered in MATH 2UU3. To align with the research questions of this project, a rubric was developed for each of the four categories of numeracy skills being assessed. These categories are as follows:

Category A: A student's ability to understand numbers (e.g., absolute size, relative size, patterns) and work with numeric information (e.g., approximate, estimate, scale, visualize)
Category B: A students' ability to use logical reasoning (such as understanding and creating cause and effect arguments) and recognize logical fallacies.
Category C: A students' ability to engage with multiple step-problems which require quantitative reasoning.
Category D: A student's ability to communicate their answers to quantitative reasoning problems and questions.

These rubrics were constructed after a preliminary literature search about numeracy and assessment of numeracy, and assessment strategies from elsewhere which we deemed to be appropriate for our study. The primary reference for this phase of research is the First Results of the Adult Literacy and Life Skills Survey published by OECD and Statistics Canada in 2013. This source outlines a large study, in which numeracy assessments were part of the concept domain. Their working definition of numeracy is as follows:

The knowledge and skills required to effectively manage and respond to the mathematical demands of diverse situations.

While this is admittedly broader than our definition, further dissection of the source identifies that the numeracy tasks demanded of participants are deeply rooted in their context, as the tasks administered to students in MATH 2UU3. Further, the document outlines what the others refer to as "numerate behaviour":

***Numerate behaviour** is observed when people manage a situation or solve a problem in a real context; it involves responding to information about* *mathematical ideas that may be represented in a range of ways; it requires the activation of a range of enabling knowledge, factors, and processes.*

Given these definitions of numeracy and numerate behaviour, the authors of this paper judge that their conception and assessment strategies are appropriate to adapt to our study.

Rubric for Category A: Understanding and Working with Numeric Information

Category Description: A student's ability to understand numbers (e.g., absolute size, relative size, patterns) and work with numeric information (e.g., approximate, estimate, scale, visualize).

	Level 1	**Level 2**	**Level 3**	**Level 4**	**Level 5**
Ability to Understand Numbers	Student skills are emerging in this trait.	Student can identify relevant numeric information but does not show interpretation in context.	Student can identify relevant numeric information from few or no distractors and shows attempt to interpret it's meaning in context.	Student can identify relevant numeric information, sometimes from other plausible distractors, and shows a valid attempt to interpret it's meaning in context.	Student can identify relevant numeric information regardless of the presence of distractors, and correctly interpret it's meaning in context.
Working with Numeric Information	Student skills are emerging in this trait.	Student attempts to employ an inappropriate quantitative reasoning skill but attempts to address the numeric information in the question with their answer.	Student attempts to employ an appropriate quantitative reasoning skill to work with presented numeric information in context, but makes three or more minor errors (e.g., calculation errors), or one major error (e.g., selecting an inappropriate computation).	Student attempts to employ an appropriate quantitative reasoning skill to work with presented numeric information in context, but perhaps makes one or two minor errors (e.g., calculation errors).	Student can employ quantitative reasoning skills to work with presented numeric information correctly within its context, and in a variety of forms.

Rubric for Category B: Logical Reasoning

Category Description: A students' ability to use logical reasoning (such as understanding and creating cause and effect arguments) and recognize logical fallacies.

	Level 1	Level 2	Level 3	Level 4	Level 5
Logical Structures	Student skills are emerging in this trait.	Student attempts to use an inappropriate choice of logical structure to evaluate the truth value of a statement and describe methods of proving or disproving a statement, OR attempts to use the correct logical structure with two or more logical errors.	Student attempts to use the correct logical structures to evaluate the truth value of a statement and describe methods of proving or disproving a statement but uses them with a minor logical error. (e.g., correctly identifying the use of the contrapositive, but misinterpreting the meaning)	Student can correctly use logical structures to evaluate the truth value of a statement and describe methods of proving or disproving a statement.	Student can clearly articulate and correctly use logical structures to evaluate the truth value of a statement and describe methods of proving or disproving a statement.

Rubric for Category C: Problem Solving using Quantitative Reasoning

Category C: A students' ability to engage with multiple step-problems which require quantitative reasoning.

	Level 1	Level 2	Level 3	Level 4	Level 5
Engaging with Multi-step Problems	Student skills are emerging in this trait.	Student selects at least one correct step in reasoning or computations to solve the problem.	Student selects some correct steps in both reasoning and computations, or most correct steps in either reasoning or computations to solve the problem.	Student selects mostly correct steps in both reasoning and computations to solve the problem.	Student selects all correct steps in reasoning and computations to solve the problem.
Computational Proficiency	Student skills are emerging in this trait.	Student is unable to construct mathematical formulae or arrive at a final solution or constructs inappropriate formulae and arrives at an unreasonable solution without addressing the issues with their solution. (e.g., having a negative number of postal codes in a combinatorial problem)	Student attempts to construct mathematical formulae and arrives at reasonable values for all aspects of problem solving or constructs inappropriate mathematical formulae and recognizes the unreasonableness of their solution. (e.g., recognizing a negative number of postal codes in a combinatorial problem is an issue)	Student constructs correct mathematical formulae and arrives at reasonable values for all aspects of problem solving or recognizes the unreasonableness of a given solution.	Student constructs correct mathematical formulae and arrives at correct values for all aspects of problem solving.

Rubric for Category D: Ability to Communicate with Quantitative Reasoning

Category D: A student's ability to communicate their answers to quantitative reasoning problems and questions.

	Level 1	Level 2	Level 3	Level 4	Level 5
Communication of Solutions	Student skills are emerging in this trait.	Student can communicate using at least one correct definition of terms and concepts, or mostly correct description of terms and concepts, in their communication of solutions.	Student can communicate using some correct definitions of terms and concepts, or mostly correct description of terms and concepts, in their communication of solutions.	Student can communicate using mostly correct definitions of terms and concepts, or mostly correct description of terms and concepts, in their communication of solutions.	Student can communicate using either correct definitions of terms and concepts, or correct description of meaning of terms and concepts, in their communication of solutions.
Explanations of Quantitative Reasoning	Student skills are emerging in this trait.	Student attempts to justify their choice of computational practices, or their quantitative reasoning practices, or why their final answer is reasonable based on the context of the question.	Student attempts to justify their choice of EITHER their computational OR quantitative reasoning practices with reference to the context of the question, and where necessary, explain why their solution is a reasonable one based on the context of the question.	Student can justify their choice of EITHER their computational OR quantitative reasoning practices with reference to the context of the question, and where necessary, explain why their solution is a reasonable one based on the context of the question.	Student can justify their choice of computational and quantitative reasoning practices with reference to the context of the question, and where necessary, explain why their solution is a reasonable one based on the context of the question.

Appendix D
ORAL CONSENT SCRIPT FOR DELAYED POST-TEST INTERVIEWS

Introduction:
Hello, I am [your name]. I am conducting interviews about numeracy skills and numeracy thinking that you have acquired through participating in MATH 2UU3, Numbers for Life. I'm conducting this as part of research, under the direction Dr. Miroslav Lovric of the Department of Mathematics and Statistics at McMaster University.

Study procedures:
I'm inviting you to do a one-on-one virtual interview that will take about 60 minutes. I will ask you a series of questions that require numeracy to interpret and solve, and to describe your problem-solving process. I will take handwritten notes to record your answers as well as use an audio recorder to make sure I don't miss what you say. If this time no longer works for you, we can set up a different time and conferencing platform that works for us both.

Risks:
It is not likely that there will be any risks to you in this study. You might worry what we will think of you, after analyzing your surveys; or you might be bothered by the conclusions we reach. If you feel uncomfortable about us using your interview for research, you have an opportunity to withdraw from this research. You will be able to withdraw from the study before, during, or any time until 48 hours after your interview by sending an email to the researcher Julie Jenkins at je.jenkins19@gmail.com. We will destroy any data you would like excluded from the study, and never use it in our research. Data collected throughout the course of research will be completely anonymized.

Benefits:
In this study, we hope to assess the success of the numeracy course Math 2UU3 to prepare students for the numeracy demands of day-to-day life. We would also like to identify instructional practices that benefit student understanding and suggest alternative instructional methodologies to improve the quality of student learning. It is our hope that we can clearly exhibit the conceptual gain in these skills and repackage some of what we've learned for pre- and in-service teachers, as well as develop similar courses for other universities. We also hope that our research, and results, will encourage faculty to consider modifying their teaching of mathematics by emphasizing student engagement in their classes.

In appreciation for your participation, we will offer you a $25 Amazon gift card. To obtain the gift card, we only require that you start the interview; we will give you a card in appreciation of your time regardless of whether or not you complete the interview.

Voluntary participation:

Your participation in this study is voluntary.

- You can decide to stop at any time, even part-way through the interview for whatever reason, or up until 48 hours after your interview.
- If you decide to stop participating, there will be no consequences to you, and you will still receive your gift card.
- If you decide to stop, we will destroy any data collected up to that point.
- If you do not want to answer some of the questions you do not have to, but you can still be in the study.

If you have any questions about this study or would like more information you can email Julie Jenkins at je.jenkins19@gmail.com
This study has been reviewed and cleared by the McMaster Research Ethics Board. If you have concerns or questions about your rights as a participant or about the way the study is conducted, you may contact:
McMaster Research Ethics Board Secretariat
Telephone: (905) 525-9140 ext. 23142
c/o Research Office for Administration, Development & Support (ROADS) E-mail: ethicsoffice@mcmaster.ca

I would be pleased to send you a short summary of the study results when I finish going over our results. Please let me know if you would like a summary and what would be the best way to get this to you.

Consent questions:

- Do you have any questions or would like any additional details? [Answer questions.]
- Do you agree to participate in this study knowing that you can withdraw at any point with no consequences to you? [If yes, begin the interview; if no, thank the participant for his/her time.]
- Do you agree that the interview can be audio recorded? [If yes, begin the interview; if no, thank the participant for his/her time.]
- Do you consent to having your interviewer take hand-written notes? [If yes, begin interview; if no, ask if they would like to begin the interview without hand-written notes.]
- Would you like to receive a summary of the study results? [If yes, ask how they would like to receive the summary. Once recorded, begin the interview; if no, begin the interview.]