Burazin A., Gula, T. & Lovric, M. (2025) Examining the effectiveness of the numeracy course 'Numbers for Life' at McMaster University. *Adults Learning Mathematics: An International Journal 19*(1), pp. 22-44

Examining the effectiveness of the numeracy course 'Numbers for Life' at McMaster University

Andrijana Burazin
University of Toronto Mississauga Canada
<a.burazin@utoronto.ca>
Taras Gula
George Brown College, Toronto, Canada
<tgula@georgebrown.ca>
Miroslav Lovric
McMaster University, Hamilton, Canada
< lovric@mcmaster.ca>

Abstract

This paper reports on our study of the effectiveness of the course instruction in the numeracy course 'Numbers for Life', offered at McMaster University, Canada, and taught by one of the co-authors. In this course students explore the ways to reason with numbers in a variety of contexts important not only for their individual lives and their community, but broader - it helps them understand the world they live in, and the challenges they will face. To answer important questions about students' learning and their development of quantitative reasoning skills, we designed a two-year research project, for which we secured government funding. We combined quantitative and qualitative data analysis methods to assess the gains in student learning and skills development using a number of instruments, including pre-test and post-test surveys, class activities, course assessments and teaching evaluations. Our research suggests that the "Numbers for Life" course instruction improves most students' numeracy knowledge and skills. We detected improvements in students' ability to understand numbers, ability to engage with logical constructions and reasoning, and ability to engage with multiple-step problems which require quantitative reasoning. The largest learning gains were detected among students who had inadequate background, as determined by their pre-test performance. We were not able to detect pre- to post-test improvements in communication (explaining what a numeric answer represents, providing a logical argument, creating a narrative about a situation involving numbers, and so on). We believe that part of the reason for this lies in the fact that students were probably not as serious, nor as patient in their approach to completing their post-test (at the end of the semester) as they were in completing a pre-test (at the start of the semester).

Key words: numeracy, effectiveness of teaching, understanding numbers, interpreting numbers, logical reasoning, communication.

Introduction

In this paper we report on the (Ontario) government-funded project whose aim was to assess the effectiveness of the instruction in the 'Numbers for Life' course offered at McMaster University, Hamilton, Ontario, Canada, in the academic years 2021/22 and 2022/23. (By 'effectiveness' we mean determining to what extent students understood and learnt the content of the course; the precise meaning is given in the section *Research Questions*).

The ‘Numbers for Life’ course was designed to address the lack of numeracy knowledge and skills among post-secondary students. Very high demand for the course suggests that, contrary to some beliefs, only teaching numeracy in primary and secondary schools does not suffice. In part, this is due to the fact that certain topics (such as mortgages, calculating interest on a credit card debt, taxes and tax brackets, consumer price index and inflation) are of little interest to primary and lower grades high school students, whereas university students find them highly relevant and important.

The topics discussed in the ‘Numbers for Life’ course include contrasting the language of mathematics and everyday language; logical reasoning, correlation and causation; addressing 'quantitative' in quantitative reasoning (e.g., numbers and related concepts, fraction, ratio and percent; absolute and relative numbers, exactness and approximation, estimation, order of magnitude and scientific notation); patterns of change; linear regression. As well, the course discusses randomness, chance and risk, intuitive understanding of basic probabilistic principles, reasoning based on standard deviations, and working with conditional probability and contingency tables. The contexts within which these topics arise include numerous situations from everyday life, human population dynamics, climate change, personal and general finance, gambling, and economy and social indicators. For a discussion and conceptualization of numeracy and numeracy tasks that informed the teaching of the course, see Gula & Lovric (2024).

By its very nature, numeracy is about problem-solving (always in authentic, real-life contexts), and classroom activities in ‘Numbers for Life’ support this view. For example, in the activity named “Wasted Water” students were asked to estimate the amount of water wasted in Canada every day by people who do not turn the water tap off while brushing their teeth. This activity is described and analysed in Burazin and Lovric (2018). Another activity asked students to calculate the interest on a credit card when its balance was not paid in full, based on the information given in the cardholder agreement.

Besides getting involved in class discussions and group work (less so, unfortunately, in virtual classes), students learn in the course by listening to lectures and reacting to provocations (used as invitations to criticize, oppose, or contradict their instructor), working through situations that require more complex numeracy reasoning (e.g., unpacking an unclear or incomplete narrative, or engaging in multi-step reasoning), and reading and analyzing narratives from various sources. As well, students are encouraged to learn on their own by working on weekly problem sets.

The main objective of our research project was to measure learning gains and skills development of students enrolled in the ‘Numbers for Life’ course, both in the short term (i.e., immediately after the course ends, compared with the start of the course), and in the longer term (one year after the students finish the course) to measure retention of knowledge and skills, which is related to the transferability of what is learned in the course. The outcomes of this research have already informed modifications and fine-tuning of the ‘Numbers for Life’ course, its curriculum, materials used, and instruction.

The population we studied consisted of two cohorts of students: those who took the ‘Numbers for Life’ course in the Fall 2021 semester (September to December 2021), and those who were in the course in the Fall 2022 semester (September to December 2022).

This project employed action research (well suited, among other goals, to inform modifications of the ‘Numbers for Life’ course curriculum and instructional interventions), quantitative, and descriptive/ qualitative research. Action research draws on well-known “‘ideas of the reflective practitioner’ and ‘reflection-on-action’: the active and critical consideration and

reflection by us, as practitioners, on such aspects as the motives behind and the consequences of our professional practice." (Kaye, 2014).

In measuring learning gains, we did not rely solely on a quantified comparison of pre- and post-test scores; instead, we probed deeper into students' work to better assess how their learning and understanding, as well as skills development, if any, changed over time. Assessing learning gains is a challenge: we cannot assume that a particular learning gain is due to a particular learning strategy; instead, we view learning as a result of the action of multiple factors, whose relative significance changes continuously throughout the learning process. The randomized comparison (treatment and control groups selected among the students in the 'Numbers for Life' course) would require differentiated instruction in the same course, which was not acceptable. Consequently, our primary strategy was to study students' responses to the survey instruments and course assessments.

We used two types of instruments:

(1) Those which compare the competencies we study (knowledge and skills) at specific temporal points, i.e., pre-test (start of the course), post-test (end of the course), and, for one cohort of students, delayed post-test (one year after they finished the course); and
(2) Those that follow a continuous development of competencies throughout the course (i.e., looking at all assessments administered in the course).

We measured the effectiveness by qualitative and quantitative comparisons of students' answers on the pre-test and the post-tests. The instruments in (1) were modified from the ones that one of the co-authors and his PhD student developed for researching a different course at McMaster University, including ensuring validity and reliability; see Appendix 1. For both (1) and (2) we used the rubrics (as is common in this situation) that follow an individual's achievement across the knowledge and skills competencies which were extracted from our research questions; see Appendix 2.

Background and Literature Review

Although there are numerous articles on numeracy and its benefits, numeracy instruction and assessment, there are very few publications that address the specifics of numeracy instruction.

The main questions we have in mind - how to teach numeracy, and how do we know that we have succeeded (or not) - are poorly covered in academic literature. Our research aims at filling this gap, by providing insights into the teaching of specific topics (see *Research Questions*) and assessing to what extent students learnt these topics. Another novelty in our research is that we assess retention by testing students one year after they completed the course. We are aware of one publication that addresses learning of numeracy in the longer term: Mandell & Klein (2009) compare two groups (those who took a specific financial literacy course, and those who did not), and conclude that there is no difference in the financial literacy behaviour between the members of the two groups at a point later in their life.

The importance of the development of numeracy knowledge and skills among students as well as adults has been recognized in research literature, as well as in various government documents (Gal, 2024; Dingwall, 2000; Dion, 2014; Geiger, Goos, & Forgasz, 2015; National Numeracy, 2019; Orpwood & Sandford Brown, 2015; see also a number of articles in Díez-Palomar et al., 2022). However, how to teach numeracy at various levels of education - from elementary, to tertiary, to adult - remains a challenge that has been addressed by (in our view, still inadequately) research and debate. Geiger, Goos, & Forgasz (2011) write: "far less is known about

how teachers [of numeracy] learn about, appropriate and then create effective mathematics teaching practices."

The beliefs that numeracy should be a by-product of traditional education (not just mathematics, but other subjects as well, as advocated by Steen (2001)) clash with the beliefs that numeracy instruction requires its own space and attention (see, for instance, the discussion by Brooks (2013) in the context of adult education). 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 campus, except to Math and Stats majors.

Among the research contributions addressing teaching of numeracy, we consulted those that were specific in studying teaching practice, and creating resources, such as Goos (2016). A number of 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.

Kaye (2014) studies teaching environments and suggests an interesting departure from teaching numbers (say, within a numeracy course) and instead focussing on topics such as data and measuring shape and space; see also Hoogland et al. (2019) and the references therein. Boaler (1993) and Hoogland & Pepin (2016) suggest approaches to creating good numeracy tasks to be used in classrooms.

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.

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 are the result of a collaboration of a number of international experts. In creating our assessment instruments (see the rubrics in Appendix 2) we were guided by the assessment frameworks for PISA and PIAAC, as well as by the 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 to the questions that helped us to answer our research questions. To judge the difficulty of our assessment tasks, we consulted Weingarten et al.'s (2018, pp 28-29) numeracy proficiency levels, as we planned most of our tasks to be at their highest level (level 4/5).

Research Questions

The research questions that guided our project are:

(R1) To what extent does the course 'Numbers for Life' develop students' numeracy skills?

Category A. How has students' ability to understand numbers (e.g., absolute and relative size, patterns) and work with numeric information (e.g., approximate, estimate, scale, visualize) changed after completing the course?

Category B. 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?

Category C. How has students' ability to engage with multiple step-problems which require quantitative reasoning changed after completing the course?

Category D. 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?

(R4) How does numeracy skills development differ between the two student cohorts?

We recognize that answering these questions in their entirety, given the vast amount of data we collected, requires lot more time than we have dedicated to this project so far. This paper is our first formal report on some findings of our research, in which we do not answer all questions, nor analyze all data that we have collected. Our forthcoming efforts will include studying the interviews we have conducted with samples of students from both cohorts, as well as answering (R3) and (R4).

Methodology

The survey population consisted of two cohorts of students in the "Math for Life" course taught at McMaster University in Fall 2021 (13 weeks, from September to December 2021) and Fall 2022 (13 weeks, from September to December 2022). About two thirds of the students gave consent for their data to be used in this research (384/592 in the Fall 2021 cohort and 304/399 in the Fall 2022 cohort), which constitutes a large sample size.

The data collected from this sample have been used to analyze and answer research questions (R1)-(R4). A group of 12 students from each cohort was recruited (through a general e-mail call to all students in both cohorts, with the selection based on the first-come-first-served principle) for follow-up individual, semi-structured, interviews. These interviews helped us to address research questions (R1), (R2), and (R3). As well, their purpose was to validate questions (through testing the survey on this subset of the population) developed to measure learning gains in categories A-D of the numeracy skills in (R1) that were not taken from other (validated) numeracy skills instruments.

The research team for this project consisted of three faculty members (authors of this paper), three graduate students and two undergraduate students.

We post-tested the Fall 2021 cohort of students again in the first week of December 2022 (we call it "retention post-test" or "delayed post-test"), to measure long(er)-term retention of knowledge and skills. Experience tells us that reaching out to students in this situation (i.e., when they are no longer part of a course) is a challenge, so we incentivized them by offering small value gift certificates for their participation. For qualitative responses, we developed four rubrics in accordance with (R1) categories A-D, to help ensure that we could effectively compare the performance of students in their pre-tests with their post-tests; see Appendix 2.

We obtained ethics clearance from 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 was fully protected.

We collected an extensive amount of data from 384 students (in the Fall 2021 cohort) and 304 students (in the Fall 2022 cohort), in order to be able to better position ourselves to draw reliable conclusions. The data collected includes:

- Course assessments (collected twice, in Fall 2021 and Fall 2022; 6 assessments in each year, collected through our learning management system)
- Pre- and post-test surveys (collected twice, in Fall 2021 and Fall 2022)
- Semi-structured interviews with a small sample of students (conducted twice, in Winter 2022 and Winter 2023, each time with 12 students)
- Retention post-test survey: Fall 2021 cohort of students was assessed a year later (in December 2022).

The pre-test, post-test and the retention post-test contained identical questions, to make it easier to compare students' answers. The students surveyed were not aware of it prior to seeing the tests.

We collected data twice (in two consecutive offerings of the course, one year apart) to generate a repeated experiment and ensure robustness. It is well known that cohorts of university students differ from year to year, and that a pedagogical intervention that leads to a success in one year might not show any significant effects the following year. By repeating the experiment, we are able to control for some of the differences between the two successive cohorts of students, which allow us to distinguish between the outcomes which are consistent and those which only occurred in one cohort.

We developed a series of test instruments by going through the following process:

- Content validity: concept domain analysis and expert consultation; several items have been taken from pre-existing instruments that we found in literature.
- Validation: through interviews with small subsets of participants we were able to evaluate whether the questions were testing the concepts and skills as intended.
- Reliability: we have repeated testing over two cohorts, except for the delayed post-test for the second cohort.

One of the largest issues we faced is related to students' attitudes towards taking the survey instruments (pre-test and post-test). For this research we assumed that, as the pre- and post-test were graded for meaningful completion (as opposed to earning course credit for accuracy and correct interpretation), some students may have lacked motivation to provide a quality answer on the post-test, since they were more familiar and comfortable with the course assessment than at the start of the course (thus, they knew how much, or how little, they had to do to earn credit). This fact renders a bare comparison between pre- and post- test scores more or less useless and was the main reason why we primarily focussed on qualitative analysis. However, as we did want to compare numeric scores, we decided to normalize the post-test scores based on the amount of time students spent completing the surveys. We are aware of alternative designs (such as having a control group of students who take a pre-test and a post-test), but that would still not address the problem.

In the next section we compare pre-test and raw (ie., unnormalized) post-test survey responses. We proceed by presenting a rationale for normalizing the post-test scores and then compare the pre-test scores with the normalized post-test scores.

Results

In this section we report on some of the results that we have obtained so far.

The nature of this research, and especially the realities of administering survey instruments (students' attitudes towards surveys, for instance) do not make it feasible to place too much emphasis on reporting results in a numeric format. However, we believe that in some cases, there is value in looking at numbers.

Several questions on the pre- and post- tests were about understanding and interpreting numbers and working with numeric information. In particular (see also Appendix 1):

- Question 1. 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.
- Question 2. [Students were given a table with nutrition facts for a container of milk (the typical "food label") which included the serving size, calories per serving, and the number of servings per container.] If you drink the entire container how many calories would you consume?
- Question 9. [Students were given a picture of a brick wall, with a woman walking alongside the wall.] 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). Explain your reasoning.
- Question 10. 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?

All four questions required submitting a narrative to explain how the answer was arrived at. To evaluate these responses, we developed a rubric (see Appendix 2) which placed each answer on a scale from level 1 (lowest) to level 5 (highest). We are aware that the rubric evaluation scale is a categorical variable, and so calculating averages (means) is not appropriate. Nevertheless, thinking of a value from 1 to 5 as an actual grade that a student would earn in the course assessment for answering the question, we can make sense of the mean value.

In analyzing student learning, we compared these numeric pre- and post-test survey scores, with the two surveys administered almost three months apart (Table 1). A body of research in mathematics education (and possibly elsewhere) uses a post-test design (with no pre-test), with the post-test administered a very short time after the intervention that is studied. If a pre-test is used, it is often administered just before the intervention, with a post-test administered right after it (hence the raw pre- to post- data is compared, with no need to normalize the scores). Consequently, we could not identify a reference in mathematics education that we could use as a guide for our situation where the pre- and post-tests were three months apart, and where we believed that the comparison of the raw (i.e., unnormalized) pre- to post-scores would not yield realistic conclusions.

Table 1 shows a comparison of the scores on the pre- and (unnormalized) post-tests for the four questions mentioned above, marked using rubrics, on a scale from 1 to 5 (Category A items listed are part of the research question (R1)). The questions in the pre-test and post-test were identical. All but one of the pre-test scores are higher than the corresponding post-test scores, indicating (on a superficial level) that the course instruction did not contribute to an improvement.

Note: the difference in the means of the two groups is not statistically significant (Welch's two-tailed t-test, p=0.4).

Table 1. Mean value of questions 1, 2, 9 & 10.

Question	Topic	Pre-test (1-5)	Post-test (1-5)
1			
	Category A: **Ability to understand numbers**	4.54	4.45
	Category A: **Working with numeric information**	4.52	4.6
2			
	Category A: **Ability to understand numbers**	4.54	4.47
	Category A: **Working with numeric information**	4.58	4.51
9			
	Category A: **Ability to understand numbers**	4.01	3.79
	Category A: **Working with numeric information**	4.27	4.11
10			
	Category A: **Ability to understand numbers**	4.46	4.44
	Category A: **Working with numeric information**	4.23	3.91

Questions 11 and 12 probed into students' understanding and the use of mathematical logic and logical reasoning (Table 2), marked using rubrics, on a scale from 1 to 5 (Categories B and D are a part of research question (R1)).

- Question 11. 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 12. 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 reinterpretation correct? Why or why not?

Again, these questions were graded on the rubric (Appendix 2). There is a significant improvement in logical reasoning (category B and category D, except in question 12) in the post-test. Unlike the topics related to numbers and developing number sense, logical reasoning is not taught (except possibly implicitly) in high school, so the effects shown here are very likely due to the 'Numbers for Life' course instruction. However, there is no statistically significant difference between the pre- and post- test means (Welch's two-tailed t- test, p=0.10).

Table 2. Mean value of questions 11 & 12.

Question	Topic	Pre-test (1-5)	Post-test (1-5)
11			
	Category B: Logical Reasoning (**Logical Structures**)	3.49	4.04
	Category D: Ability to Communicate (**Communication of Solutions**)	3.45	4.54
12			
	Category B: Logical Reasoning (**Logical Structures**)	3.32	3.67
	Category D: Ability to Communicate (**Communication of Solutions**)	3.24	3.17

In Table 3 we compare the performance of the Fall 2021 cohort on five multiple choice and true/false questions (questions 3-7) on the surveys administered during the course (pre-test in September 2021, post-test in December 2021) and one year later (retention (delayed) post-test in December 2022). In this case, the questions were marked as correct or incorrect only; the fractions in the table represent the ratio of correct answers to the total number of answers to a particular question. (See Appendix 1 for the statements of these questions.)

Table 3. Students' performance on questions 3-7.

Question	Topic	Pre-test	Post-test	Delayed post-test
3	Reasoning about fractions: how the change in the numerator affects the value of a fraction	376/384 (97.92%)	354/374 (94.66%)	107/119 (89.92%)
4	Comparing chance events when probabilities are given in the form "a in b" (as in 2 in 100)	338/384 (88.02%)	329/374 (89.97%)	88/119 (73.95%)
5	Reasoning about the relative change when data is presented in a bar chart	165/384 (42.97%)	202/374 (54.01%)	64/119 (53.78%)
6	Distinguishing between correlation and causation in the case of disease and symptoms	225/384 (58.59%)	249/374 (66.58%)	65/119 (54.62%)
7	Venn diagram representation of a relationship between two populations	349/384 (90.89%)	337/374 (90.11%)	103/119 (86.55%)

The questions on the three instruments were identical, except that we changed questions 3 and 7 from multiple choice (on the pre-test and post-test) to a question (on the delayed post-test) where students had to make a choice and justify it ("explain" type of a question). (For the analysis in Table 3 we only looked at the student's answers to these questions and not at their rationale.) By asking for a justification, we obtained data which are more revealing in showing how students understood the concepts and articulated their answers. For instance, consider

- Question 3. 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. What is the relationship between Person A's and Person B's energy demands? [Added to the delayed post-test: Show your reasoning.]

Students' responses show different approaches (which also shows that analyzing answers qualitatively is, in our case, superior to a numerical analysis):

- Student 11039 selects relevant data and ignores the rest: "Since the heart rates are the same you can take that out of the equation. Thus, to compare the energy needed the cholesterol levels can be looked at, since person A has a lower cholesterol level they require less energy in comparison to person B."
- Student 36322 correctly uses the terms "proportional" and "inversely related": "Since Person A's cholesterol is lower, Person A will have a lower energy demand than Person B. This is because cholesterol and energy demand are proportionally related, whereas energy demand

and heart rate are inversely related. As well, since heart rate is the same and 9.5 is a constant, the only variable that can affect their energy demands is cholesterol."

- Student 76331 reasons about how the value of the numerator affects the fraction: "Person A would have a lower energy demand than Person B. This is because Person A has a smaller number as the numerator, which would make for a smaller number overall if the denominator stays the same. Since they have the same heart rate, and the numerators are being multiplied by the same number, the only thing changing is the numerator itself, which is the cholesterol level."
- Student 33494 shows a "mathematical" approach, using a simple example to illustrate the point: "Energy demand = (cholesterol level*9.5)/heart rate A - lower cholesterol --> lower energy demand e.g. A=1, heart rate=1 --> ED = 9.5 B=2, heart rate=1 --> ED=19"

Questions 3, 4, and 7 (Table 3) do not appear to show significant differences between pre-tests and post-tests; in both cases the correct response rates are very high. Questions 5 and 6 show improvement in the post-test. As expected, the delayed post-test rates suffered from forgetting due to the temporal distance, but definitely do not follow a typical, concave upward, sharply declining forgetting curve (the classical Ebbinghaus curve).

Normalizing post-test scores

Looking at the results presented in the previous section, we felt that it did not represent the reality that we observed in the classroom. Students seemed to be learning, but that did not surface in the results. In order to produce a more meaningful numeric comparison of pre-test and post-test answers, we decided to normalize the data. The main reason for doing so came from our examination of students' answers, as we realized that the same student answered the same question better (more thoroughly, with adequate correct narrative explanation) in the pre-test that in the post-test. Here is a typical example: student 5479's answer to Question 1 (stated above): "less than S dollars, since the tax is applied to the newer, lesser price. The discount is applied first then the tax. even though they are the same percentage since they are applied to different prices they will have different absolute value. The tax will be applied to a lower price and the discount is reduced from a larger price." (pre-test); "Less than s dollars, the tax is added after the discounted price so the percentage is different" (post-test).

Both answers are correct, but it is clear that the student put more thought and effort into answering the question on the pre-test. It is clear from the post-test that the student had a correct answer in mind, but did not communicate it as thoroughly as in their pre-test answer.

Another factor that, very likely, negatively affected the quality of students' responses to the post-test is a general fatigue (as well as stress and a decline in motivation) that students experience near the end of a semester.

The quality of responses (the "score") to Question 1 (as well as many other questions) was judged based on the rubrics that we developed (see Appendix 2; question 1 was judged based on the first rubric shown there). The average scores for Question 1 on a pre-test and post-test (Category A, first line in Table 1) were 4.54 and 4.45 respectively. To interpret these numbers would mean to claim that, on average, the students did better (or about the same, as the two averages are close to each other) on the pre-test than on the post-test. However, as the example above shows (and many others), the student did know the answer, did know the correct explanation, it was just that their post-test communication of the answer lacked the clarity and detail of their pre-test response.

As mentioned earlier, because of the way the surveys were administered (students earning full credit for a meaningful completion, rather than earning points for correctness of their answers), and the students learning as the semester advanced that they can earn full credit by providing only partial answers, the post-test responses were often incomplete in terms of narrative explanations. (Assigning a numeric grade was deemed inappropriate, as some questions on the pre-test included content that was not yet covered in the course.)

Not finding any guidance in literature, and trying different approaches (normalizations), we decided to adjust the post-test scores based on the average amount of time that students spent on the pre-test compared to the post-test as a proxy. This is far from perfect: it could be that the students knew (or learnt) the material, and were able to answer questions more quickly on a post-test than on a pre-test.

The pre-test and the post-test were administered online, via the learning management system, and students were allowed up to 24 hours to complete them. The mean completion time on the pre-test was 137.51 minutes (with the range from 5 minutes to almost 24 hours); the mean completion time on the post-test was 116.89 minutes (with the range from 1 minute to almost 24 hours). The ratio of the two is 137.51/116.89=1.176, telling us that, on average, a student spent 17.6% more time on the pre-test than on the post-test. As a normalizing factor, to account for the effects of students' attitudes on the post-test, we decided to use one half of this ratio, i.e., 8.8% or the factor of 1.088, and obtained the numbers in the last columns of Tables 4 and 5. We took one-half, as we believed that using the full ratio would overamplify the numbers, resulting in very high normalized post-test scores. This decision can also be interpreted as an indication of our hesitation about the appropriateness of this normalization. As we did not find anything in literature that would support, or not, this approach, and as we believe that it is somewhat useful, we decided to use it. Table 4 shows the mean values on questions about understanding numbers and working with numeric information, marked using rubrics, on a scale from 1 to 5, with both the raw (unnormalized; shown in grey) and normalized post-test means for comparison (Category A items listed are part of the research question (R1).)

Table 4. Unnormalized and normalized mean values of questions 1, 2, 9 & 10

Question	Topic	Pre-test (1-5)	Post-test (1-5)	Normalized Post-test
1				
	Category A: **Ability to understand numbers**	4.54	4.45	4.84
	Category A: **Working with numeric information**	4.52	4.6	5.00
2				
	Category A: **Ability to understand numbers**	4.54	4.47	4.86
	Category A: **Working with numeric information**	4.58	4.51	4.91
9				
	Category A: **Ability to understand numbers**	4.01	3.79	4.12
	Category A: **Working with numeric information**	4.27	4.11	4.47
10				
	Category A: **Ability to understand numbers**	4.46	4.44	4.83
	Category A: **Working with numeric information**	4.23	3.91	4.25

With this normalization, we see an improvement in the post-test performance compared to the pre-test (Table 4). The normalized post-test scores in Table 5 amplify the increases in the unnormalized post-test scores (shown in grey). Categories B and D are a part of research question (R1).

Table 5. Unnormalized and normalized mean values of questions 11 & 12.

Question	Topic	Pre-test (1-5)	Post-test (1-5)	Normalized Post-test
11				
	Category B: Logical Reasoning (**Logical Structures**)	3.49	4.04	4.40
	Category D: Ability to Communicate (**Communication of Solutions**)	3.45	4.54	4.94
12				
	Category B: Logical Reasoning (**Logical Structures**)	3.32	3.67	3.99
	Category D: Ability to Communicate (**Communication of Solutions**)	3.24	3.17	3.45

Discussion

Our analysis (keeping in mind the normalization that we applied) suggests that, overall, students taking the course 'Numbers for Life' improved their numeracy knowledge and skills. Of course, such a general statement as an answer to our research questions about something as complex as learning requires a closer look.

For background knowledge, numeracy relies on high school number sense, arithmetic, some basic geometry (all of which are taught in grade 10 and/or earlier), and logic (although not taught explicitly in high school, it is sometimes used without much attention paid to it). It turns out that students whose mathematical background is poorer (using low score on the pre-test as a proxy), learn more in the course (as judged by comparing their post-tests to pre-tests and by the follow-up interviews) than those who showed high(er) achievement on the pre-test. We now illustrate this finding using a student's answer the following question:

- Question 13. 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.

We compare student 71217's pre- and post-test responses:

[pre-test] "(7 x 26)^2 x (8 x 3^10)^4 = 1.65e27 combinations are possible."

[post-test] "L= 7 L=26 D= 8 D=10 D=10 D=10; Maximum number of postal codes= 7*26*8*10*10*10= 1 456 000 possible combinations."

The pre-test answer shows minimal evidence of the student dissecting the problem into smaller parts or of them interpreting the statement of the problem. It shows a misunderstanding of the principles of counting, and the lack of critical reflection about the (impossible) size of their final answer. Their post-test response shows significant improvement, as the student unpacked each character in the postal code in terms of the number of possible options, and used those options to successfully apply the multiplication principle and arrive at a correct final value. We note that there was little (pre- to post-) change among high-achieving students on this question ("high-achieving" qualification is based on the pre-tests scores).

Question 10 (ability to communicate answers to quantitative reasoning and ability to understand numbers; the statement appears earlier in this paper and also in Appendix 1) asks students to calculate and identify the types of change as relative or absolute. Students' ability to

communicate showed varied improvement as a result of taking the course. The answers of the students who improved typically did not change in content but increased in detail. For instance, student 58000's pre-test response is: "The relative change is different," compared to their post-test response: "Relative growth of X=(200-100)/100 =100%, relative growth of Y=(1100-1000)/1000=10%; although they have the same absolute growth, island X has 100% relative growth while island Y only has 10% relative growth." The key difference in responses is the added technical skill of using a relative change formula, adding language such as "absolute growth" and "relative growth", and the use of percentage change to interpret key differences between the population changes.

In the pre-test responses, we often saw that students found it challenging to explain their conceptual thinking or mathematical process. They are often correct in calculating rates of change but have yet to explain how they got the answer and their conceptual thinking behind it. One of the reasons students need help distinguishing between relative and absolute change is that the formulas for calculating them are very similar, leading to confusion. Another reason could be that students need better foundation in understanding basic concepts such as fractions, decimals, and percentages which, as we can see, are necessary for understanding the difference between relative and absolute change. On the other hand, while examining the post-test responses, it was evident that there was an improvement in student's ability to explain their thinking and they were able to explain clearly how they got to the answer.

In quantitative terms: 70 students received a lower rating (between 0-3, out of 5) on Question 10 on the pre-test. In many cases, their calculations were correct, but the interpretations were not there, or were vague or incorrect. On the post-test, this number dropped to 44. In our view, this shows a considerable change in students' understanding and conceptual thinking. This drop could be attributed to the interventions in class, such as covering the relevant topics in detail, providing additional practice problems, and interactively presenting the material. As well, this decrease is due to students being given more opportunities to develop critical thinking and problem-solving skills in topics not directly related to the content of the question.

Most responses showed little change between pre- and post-test from the communication perspective; if students answered using detailed descriptions or calculations in their pre-test answer, most gave a similar (sometimes lower) quality of response by the end of the course in the post-test. It is possible that this is due to the students' attitudes towards the assessment (as commented earlier). Other items intended to measure this outcome (ability to communicate answers to quantitative reasoning) yield similar responses, in that students developed appropriate language to describe their ideas and interpret the items, but the meaningful pre- to post- change in the quality of responses was limited.

The shift to using more formal narratives and logically correct explanations in the post-test was observed in Question 12 as well (recall that the question asks: "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?").

Student 13970 writes in their pre-test: "Yes, this reinterpretation is a correct conclusion to make (even if the actual statements are not entirely true). This is because the first statement assumes that not studying causes one to not pass the test. If it is assumed that there is a causation between the two variables, then the opposite would also be true (i.e., studying causes one to pass the test)." Their post-test reply, although not complete, shows a correct shift toward formalism: "No, these statements are converse of each other, which does not always deliver the same conclusion (ie, A implies B, but B does not necessarily have to imply A). If they were contrapositive, they would deliver the same conclusion."

As part of the assessment of the retention of numeracy knowledge and skills (research question (R2)), we looked into interview transcripts and students' responses to the in-class assessment question "What will you remember from this course five years from now?"

An overwhelming majority (85-90%) of students stated that they will remember concepts and facts, as well as valuable pieces of advice from the lectures on financial literacy, such as mortgages, credit cards interest, taxes and tax brackets ("I did not know that different parts of your income are taxed at different rates"), understanding debt and interest ("it is useful to open a line of credit when borrowing money as the interest on a line of credit is lower than what is charged on a credit card"), consumer price index and inflation (See Mundy (2011) for a discussion of financial capability, and the ways it can be improved). That students are so concerned about mortgages is very likely a reflection of the media frenzy about Gen Z people not being able to afford to buy a home (Sodha, 2022; Carrick, 2022). One student said that they saved the mortgage calculator we used in class on their laptop for later use.

In terms of impact, the following longer response summarizes what numerous students said: "This course has taught me several life lessons on numerical literacy which I will take with me and hopefully be able to share with my children one day [… Instructor said] that if you could not pay it by the end of the month, it should not be on your credit card. This line helped me reflect on my spending habits and I always think about this when making any purchase on my credit card now."

Although this is good news, it remains to be seen how it will play out in students' lives. The study by Mandell & Klein (2009) is not optimistic: the authors studied a course on personal financial management, and concluded that "those who took the course were no more financially literate than those who had not. In addition, those who took the course did not evaluate themselves to be more savings-oriented and did not appear to have better financial behaviour than those who had not taken the course" (p. 15). On the other hand, Sunderaraman, Barker, Chapman, & Cosentino (2022) claim that -"there is evidence that numeracy is one of the strongest predictors of financial decision making" (p. 711).

Following financial literacy, the most quoted items that students claim they will remember are: always reading the small print (especially when facing important documents such as loans, mortgages and car rental agreements); importance of understanding how the same numeric information can be framed (i.e., interpreted) in variety of ways; understanding numbers (size, absolute/relative, approximations, estimations), and communicating probability and risk.

Conclusion

Our conclusion (evidenced through comparing the gains in learning and skills from the pre-test to the post-test, and with the normalization that we introduced) suggests that the numeracy instruction in the university-level course 'Numbers for Life' offered at McMaster University improves, most students' numeracy reasoning, knowledge and skills. (We emphasize that in our research we focussed on the content and skills development, rather than on the pedagogical approaches to teaching the material in the course.)

This improvement is most pronounced among the students who came into the course with lower levels of numeracy skills, as suggested in our pre-test, administered early in the semester. (In our forthcoming publication we will present and discuss evidence for this claim.) This weak preparation consists of challenges in working with arithmetic concepts (such as comparing fractions or counting), mathematical logic (which is, if at all, covered only implicitly in high

school), and communication of mathematical results and ideas (including checking answers to see if they make sense within the context of a question).

In our view, this finding suggests a possible modification of teaching mathematics content in high school (and elementary grades as well!), by focussing more on meaningful numeracy tasks and learning, for instance, about numbers and number sense by working on such tasks. Evidence suggests that working on meaningful problems (open ended or not, truly authentic and real-life or not) enhances learning mathematics at all levels (Hendriana, Johanto, & Sumarmo, 2018; Nunokawa, 2005; Căprioara, 2015).

Our conclusion that students with weak backgrounds learned more suggests that with appropriate content and approach we can bring these students up to speed in their learning. This might be good news: by focussing on thinking with numbers, we can make mathematics more attractive to traditionally under-represented majors, such as humanities and social sciences. By replacing traditional mathematics content (such as calculus) with numeracy for these students, we have confidence that they will not only learn more math, but once they get through the numeracy course, they might feel motivated and encouraged to take more math courses.

We were able to detect improvements in students' ability to understand numbers, ability to use logical reasoning and ability to engage with multiple-step problems which require quantitative reasoning. One skills area where we were not able to detect improvement is communication (explaining what a numeric answer represents, providing a logical argument, creating a narrative about a situation involving numbers, and so on).

Another important aspect of learning numeracy is the awareness of framing, i.e., the awareness of the ways in which numeric information is, or can be, presented. Replying to "What will you remember five years from now," students said they will remember to read "fine print," especially when having to sign a financial document, such as a loan or a mortgage. The understanding that in real life numbers never come in their abstract form but always in context (and thus require particular attention) is a truly valuable lesson. This is what distinguishes a word problem in mathematics (as done in calculus, for instance) from a numeracy task (as done in the 'Numbers for Life' course).

In terms of retention, the majority of students said that they will remember some concepts from financial mathematics, due to their importance in their present life, and in the future. We did not focus our attention on analyzing the frequencies of specific topics, or the details in student's responses, but probing this further would definitely be worthwhile.

As we were not able to locate a reference about research that assesses learning of numeracy within a course and over a longer time period, we believe that our work is a novel contribution to the body of research in this area. Our efforts are now focussed on answering the research questions we did not discuss in this paper, on substantiating certain claims that we made in this paper (but did not provide solid evidence for) and on investigating other possible ways to normalize the post-test scores, to allow for a meaningful (i.e., more generally acceptable) comparison with the pre-test scores.

The limitations and weaknesses of our research are typical for the kind of investigation we have done and the methodology we employed. For instance, the sample (i.e., the students in the two cohorts who gave consent to use their data in research) might be biased, as possibly better students are more likely to participate. The group of students who agreed to be surveyed in the retention/delayed post-test could be biased towards those who achieved higher grade in the course.

Another issue (again, typical for all surveys) is the "survey fatigue," meaning that the students did not answer the questions near the end of a survey instrument as carefully and thoughtfully as those at the start of the survey. In comparing students' pre- and post- tests, we introduced a normalization (based on the times students used to complete their tests), which, as it has not been justified by literature, could limit the validity of our claims.

A possible improvement to survey design, when replicating the study would be to add a grading component to the surveys, which might encourage students to give their best effort on written-response items.

As mentioned earlier, this project is not over. We hope that our further analysis will produce refinements of some conclusions presented here, as well as new results and insights into students' learning of numeracy. Once we complete our analysis, we will go back to our surveys and methodology (e.g. considering what other ways of comparing students' gains in knowledge and skills would be appropriate) to see how they can be strengthened - for a possible continuation of this project, as well as for related projects (such as probing into effectiveness of other courses we teach). Our desire is that similar research into the effectiveness of teaching numeracy is conducted elsewhere, as the results of such a repeated experiment would strengthen some of our conclusions, or, on the other hand, question their strength.

Acknowledgements

We are grateful to the Higher Education Quality Council of Ontario (HEQCO) for their interest in our project, and their financial support. We thank our graduate students [names removed] for their enthusiasm, hard work, and dedication to this project, and the reviewers, whose comments helped us to look at our research with different eyes, which contributed to the improvements we made.

References

Boaler, J. (1993). The role of contexts in the mathematics classroom: Do they make mathematics more "real"? *For the Learning of Mathematics*, *13*(2), 12-17.

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

Burazin, A. & Lovric, M. (2018). Rethinking teaching first-year mathematics in university. In J. Holm & S. Mathieu-Saucy (Eds.), *Proceedings of the 2018 Annual Meeting of the Canadian Mathematics Education Study Group /Groupe Canadien d'Étude en Didactique des Mathématiques*. CMESG/GCEDM, pp 123-130. Available at: https://www.cmesg.org/past-proceedings/ .

Căprioară, D. (2015). Problem Solving - Purpose and means of learning mathematics in school. *Procedia - Social and Behavioral Sciences*, *191*, 1859-1864.

Carrick, R. (2022). A gen Z on home ownership: 'I just don't want to put myself through that'. *The Globe and Mail.* Retrieved 15 January 2024, from https://www.theglobeandmail.com/investing/personal-finance/household-finances/article-gen-z-millennial-home-ownership-mortgages/

Díez-Palomar, J., Kelly, B. & Jarlskog, L.(Eds.) (2022). Numeracy and social impact. *Proceedings of the 29th International Conference of Adults Learning Mathematics: A Research Forum (ALM).* Barcelona, Catalonia. Available at: https://alm-online.net/alm29_conference/

Dingwall, J. (2000). *Improving Numeracy in Canada.* National Literacy Secretariat. Ottawa: Human Resources and Social Development Canada. Retrieved 15 January 2024, from http://publications.gc.ca/collections/collection_2008/hrsdc-rhdsc/HS38-13-2000E.pdf.

Dion, N. (2014). Emphasizing Numeracy as an Essential Skill. Toronto: Higher Education Quality Council of Ontario. Retrieved 15 January 2024, from https://www.heqco.ca/SiteCollectionDocuments/Numeracy%20ENG.pdf.

Gal, I. (2024). Adult education in mathematics and numeracy: a scoping review of recent research. *ZDM Mathematics Education 56*, 293–305 (2024). https://doi.org/10.1007/s11858-024-01549-z

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.

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.

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

Gravemeijer, K. (1997). Solving word problems: A case of modelling? *Learning and Instruction*, *1*(4), 389-397.

Gula, T. & Lovric, M. (2024). Numeracy Tasks: Inspiring Transfer Between Concrete and Abstract Thinking Spaces. Numeracy *17* (1). DOI: https://doi.org/10.5038/1936-4660.17.1.1447

Hendriana, H., Johanto T., & Sumarmo, U (2018). The role of problem-based learning to improve students' mathematical problem-solving ability and self confidence. *Journal on Mathematics Education*, *9*(2), 291-300.

Hoogland, K. & Pepin, B. (2016). The intricacies of assessing numeracy: Investigating alternatives to word problems. *Adults Learning Mathematics: An International Journal*, *11*(2), 14-26.

Hoogland, K., Auer, M., Díez-Palomar, J., O'Meara, N., & van Groenestijn, M. (2019). Initiating a common European Numeracy Framework. *Eleventh Congress of the European Society for Research in Mathematics Education, Utrecht University*, Feb 2019, Utrecht, Netherlands. hal-02409293

Karaali, Gizem (2008). Word problems: Reflections on embedding quantitative literacy in a calculus course. *Numeracy*, *1*(2).

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

Mandell, L. & Klein, L. S. (2009). The impact of financial literacy education on subsequent financial behavior. *Journal of Financial Counseling and Planning*, *20*(1).

Mundy, S. (2011). Financial Capability: Why is it important and how can it be improved? Perspective Report. CfBT Education Trust. UK. Retrieved 3 August 2024, https://financialliteracygroup.com/wp-content/uploads/2015/10/SM3-Improving-Financial-Literacy.pdf

MyWellesley (nd). *Study Packet for the Quantitative Reasoning Assessment*, Wellesley College. Retrieved 15 January 2024, from https://www.wellesley.edu/qr/qr-assessment-study-packet

National Numeracy (2019). *Why is Numeracy Important?* Retrieved 15 January 2024, https://www.nationalnumeracy.org.uk/why-numeracy-important.

Nunokawa, K. (2005). Mathematical problem solving and learning mathematics: What we expect students to obtain. *The Journal of Mathematical Behavior*, *24*(3–4), 325-340.

Orpwood, G., & Sandford Brown, E. (2015). *Closing the Numeracy Gap: An Urgent Assignment for Ontario, October 2015.* Retrieved 15 January 2024, http://www.numeracygap.ca/assets/img/Closing_the_numeracy_gap_V4.pdf.

Sodha, S. (2022). Generation Z faces a brutal truth. Finding a decent home will be a lifelong struggle. *The Guardian.* Retrieved 15 January 2024, from https://www.theguardian.com/commentisfree/2022/aug/28/life-is-better-now-than-in-the-1950s-but-at-least-that-generation-could-afford-a-home

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.

Sunderaraman P., Barker M.S., Chapman S., & Cosentino, S (2020). Assessing numerical reasoning provides insight into financial literacy. *Applied Neuropsychology; Adult* 29(4), 710-717.. https://doi.org/10.1080/23279095.2020.1805745

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.

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. Retrieved 15 January 2024, from

https://www.heqco.ca/SiteCollectionDocuments/FIXED_English_Formatted_EASI%20Final%20Report%282%29.pdf.

Appendix 1: Pre-Test, Post-Test and Delayed Post-Test Survey 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 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?

Question 3

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.

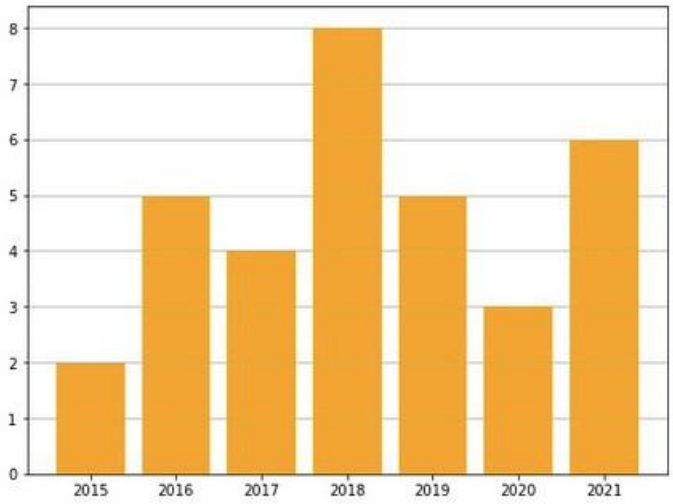
Question 4

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 5

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.

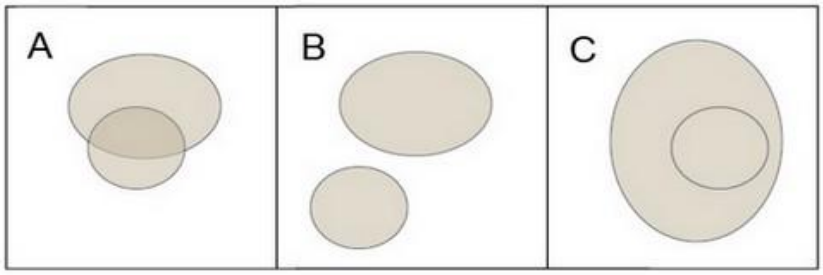
Question 6

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 7

"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 8

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?

Question 9

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.

Question 10

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?

Question 11

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 12

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 13

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 2: Grading Rubrics for Narrative Responses

Rubric for Category A: Understanding and Working with Numbers and Numeric Information

	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 its meaning in context.	Student can identify relevant numeric information regardless of the presence of distractors, and correctly interpret its 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: Using Logical Reasoning and Recognizing Logical Fallacies

	Level 1	Level 2	Level 3	Level 4	Level 5
Correlation/Causation and **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: Multiple Step Problem Solving Using 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 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 problem solving or recognizes the unreasonableness of a given solution.	Student constructs correct mathematical formulae and arrives at correct solving.

Rubric for Category D: Ability to Communicate Answers to Problems with Quantitative Reasoning

	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.