Rochester Institute of Technology

RIT Digital Institutional Repository

12-19-2013

Discovering real-world usage scenarios for a multimodal math search interface

Keita Del Valle Wangari

Recommended Citation

Discovering real-world usage scenarios for a multimodal math search interface

by

Keita Del Valle Wangari

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Human-Computer Interaction

Rochester Institute of Technology

B. Thomas Golisano College
of
Computing and Information Sciences

Department of Information Sciences and Technologies

December 19, 2013

Rochester Institute of Technology

B. Thomas Golisano College
of
Computing and Information Sciences

Master of Science in Human-Computer Interaction

Thesis Approval Form

Student Name: Keita Del Valle Wangari

Thesis Title: Discovering real-world usage scenarios for a multi-modal math search interface

Thesis Committee

Name	Signature	Date
Richard Zanibbi		
Chair		
Michael Yacci		
Committee Member		
Evelyn Rozanski		
Committee Member		

Acknowledgements

Abstract

To use math expressions in search, current search engines require knowing expression names or using a structure editor or string encoding (e.g., LaTeX) to enter expressions. This is unfortunate for people who are not math experts, as this can lead to an "intention gap" between the math query they wish to express, and what the interface will allow. m_{in} is a search interface that supports drawing expressions on a canvas using a mouse/touch, keyboard and images. We designed a user study to examine how the multimodal interface of m_{in} changes search behavior for mathematical non-experts, and discover real-world usage scenarios. Participants demonstrated increased use of math expressions in queries when using m_{in}. There was little difference in task success reported by participants using m_{in} vs. text-based search, but the majority of participants appreciated the multimodal input, and identified real-world scenarios in which they would like to use systems like m_{in}.

Table of Contents

1. Introduction

This is a first study reporting on user experiences with a query-by-expression interface supporting handwritten input. Our study was designed to gain insight into whether expressions would be desirable and/or useful to non-experts (in our case, college students) when conducting math-related searches. We reasoned that the visual appearance of an expression may be easier and more natural for non-experts to represent directly when searching for mathematical information. To explore this, participants were allowed to use a prototype of a new multimodal query-by-expression interface called m_{in}. In addition to observing use of the new interface, a secondary purpose of the study was to observe conventional search methods (including books and class notes) so that math search behaviors that are tool-independent could be identified. More specifically, the goals of this study were: 1) to observe whether m_{in} changes user search behavior, and 2) document relevant use cases for math search interfaces by mathematical non-experts. The study achieved these goals, making a number of observations as to how m_{in} impacted the use of expressions in search queries and affected search time, and discovering cases where non-expert math students felt query-by-expression would have been useful in the past, as well as in the future.

First, in Chapter 2, motivation for the study is presented by discussing relevant research into the visual perception of math and prior work regarding math input. Included is a survey of the current landscape of math search interfaces, followed by a discussion of their usability issues and an introduction to the m_{in} interface. The study design is then presented in Chapter 3 – significant effort went into designing the study tasks so that they may provide a model that may be used and adapted by researchers in future studies. For example, the use of task scenarios where participants find information to teach a colleague about a topic, as opposed to finding it for themselves, helps allay common math anxieties for participants while still asking them to locate math-related information. Chapter 4 presents the study results along with a discussion regarding their meaning, impact and contribution to the HCI community. Finally, in Chapter 5, key findings are summarized and ideas for future studies are presented. Key contributions of this study are 1) the conclusion that non-experts in the math domain may not be served well with current math search interfaces and that query-by-expression may be the missing piece that closes the intention gap in interfaces that allow only sentential or template math input, 2) the

observations of positive and negative ways in which a multimodal query-by-expression interface can change search behavior, and 3) the discovery that there are indeed real world use cases for such technology.

2. Background

Key objectives of this study are to understand the behavior of the non-expert math searcher and to inform the future development of math search interfaces so that the challenges of non-experts in the math community are considered and accounted for in search interface design. To do this, it's important to first understand why one might be motivated to believe there are any interface issues to consider with this particular audience. So, in this section, the motivation for pursuing this study is outlined. It highlights some of the more relevant research done in the areas of visual math perception and math input, surveys the current state of math search interfaces, describes the usability shortcomings of those interfaces with regard to a non-expert math community, introduces a new multimodal math search interface, and provides support for the methodology we use to evaluate the new interface.

2.1 Perceiving Math

Math notation was designed to encourage us to think visually (Anthony, Yang & Koedinger, 2005). In fact, research has been done that shows the appearance of a math equation, particularly algebraic expressions, affects our reasoning about that equation even when we know it shouldn't. Landy & Goldstone ran a series of four experiments where participants in each experiment were asked to judge whether given equations were valid or not. The equations were structured with either nonstandard spatial relationships, irrelevant or induced groupings, or manipulated spacing and those factors influenced the participants' problem-solving, even when they were given feedback during the experiment and even when they recognized that they were likely being misled by their perceptions (Landy & Goldstone, 2007b). This contradicts the long-held view that mathematical reasoning is purely symbolic. In other experiments run by Landy & Goldstone, participants were presented with natural language word equations and asked to either handwrite the equation using formal math notation or use a keyboard to enter the formal equation into an interactive website. The results showed that participants would systematically space their equations even when it was unnecessary (Landy & Goldstone, 2007a). Landy & Goldstone suggest that the whole symbolic system itself may be partially grounded in visual structure since people seem to be cognitively affected by even small changes in the spatial arrangement of an expression (Landy & Goldstone, 2007a).

Lending more credence to the importance of the physical layout of math equations is the fact that math notation is now being considered to be not just sentential but diagrammatic, as well (Landy & Goldstone, 2007a). With diagrams, it has been shown that people draw conclusions from the visual detail of the diagram as well as from the information being presented in the diagram. This is in keeping with research findings that the degree to which mathematical rules are applied are affected by visual perception, leading Landy and Goldstone to conclude that "formally symbolic reasoning is more visual than is usually proposed" and that there exists "a genuine cognitive illusion in the domain of mathematics" (Landy & Goldstone, 2007b). The cognitive illusion being that the physical spacing, the form of the notation, and the familiarity of variable names changes how one reasons about a mathematical expression.

Landy & Goldstone's research has shown that the degree of these effects may vary based on the past experiences of the person, e.g., what they learned from their teachers, what they learned from their parents, etc. We further reason that the degree to which the visual impacts the logical may be higher in math novices than in math experts. In fact, Landy & Goldstone feel it would be helpful if greater thought were given to the physical presentation of math to students. They point out that even when they were just writing out math equations without solving them, their test subjects still relied on the same spacing practices they used when actually problem-solving which possibly explains why they're so sensitive to syntactic form. They further point out that some mathematical notations, such as subscripts, serve no mathematical function at all – their role is strictly that of a cognitively helpful device. (Landy & Goldstone, 2007a)

In the literature, typed input has been found not optimal for representing mathematical notation (Anthony, Yang & Koedinger, 2005) and equation editors fall short as hunting for appropriate symbols is tedious and equation layout changes are difficult (Smithies, Novins & Arvo, 2001). Users were asked to use and rate keyboard-and-mouse, handwriting-only, speech-only, and handwriting-plus-speech and found handwriting-only to be the most natural and satisfactory (Anthony, Yang & Koedinger, 2005). Later, those results were found to generalize to middle- and high-schoolers and to simpler, easily typed equations (Anthony, Yang & Koedinger, 2007).

However, a prior research study on math information retrieval indicated that participants could not picture a scenario where entering a written math expression as a search term would be useful, doubted the value of such query-by-expression capabilities, and said they would prefer to input LaTeX due to their familiarity with it (Zhao, Kan, & Theng, 2008). Moreover, the researchers in the Zhao study stated "it is clear that users find text input the most viable form of searching" and that participants found "that specialized input modalities for equations are unwieldy." But that study did not focus on mathematical non-experts and participants were not given a prototype for hands-on evaluation. We question whether conventional, text-based keyword search alone is the best method for expressing the query need of a mathematical non-expert.

Based on the research findings on the visual aspect of math reasoning to the domain of math search, it seems there may be room for improvement in the search interfaces currently available. The conventional text keyword entry box found on standard search interfaces only supports a sentential form of math input. There is no affordance for the diagrammatic aspects of math notation. Since many math equations are learned, consumed and recalled visually, it is possible that query-by-expression is a more natural way for non-experts to express a math search need (Zanibbi & Blostein, 2012), and thus could change the user's search behavior in ways that positively impact their search experience. For example, this may lead to longer search queries that lead to more relevant results.

In the search domain, searching for math-related information can involve words and math expressions, so a multimodal interface that supports the two-dimensional, spatial aspect of a complex mathematical expression, while at the same time still supports linear text input for keywords, might be preferred by non-experts in math.

2.2 Existing Math Search Interfaces

Below is a brief survey of current math search interfaces along with interface screenshots, descriptions of their intended use and the types of input they accept. The thirteen math search interfaces presented in this section demonstrate the math entry modalities more commonly

available for math searchers. The first eight interfaces discussed below require the searcher to know LaTeX, TeX, MathML or a similar encoding language and they demonstrate the sentential form of input discussed earlier where support for the diagrammatic aspects of complex math expressions is missing.

The next four interfaces require the use of a template-based equation editor to build the math expression. While equation editors do lend support for the two-dimensional, spatial aspect of complex math expressions, building and modifying the expression can be tedious, as discussed earlier. Lastly, the browser extension discussed below is not technically a search interface on its own but offers a method by which the user can submit a math expression for search directly from where it is being displayed on a web site. This has its limitations in that the number of sites it works on is limited and, of course, the user is limited to submitting an initial search for the expression exactly as it's displayed on the web site.

Springer LaTeX[1]
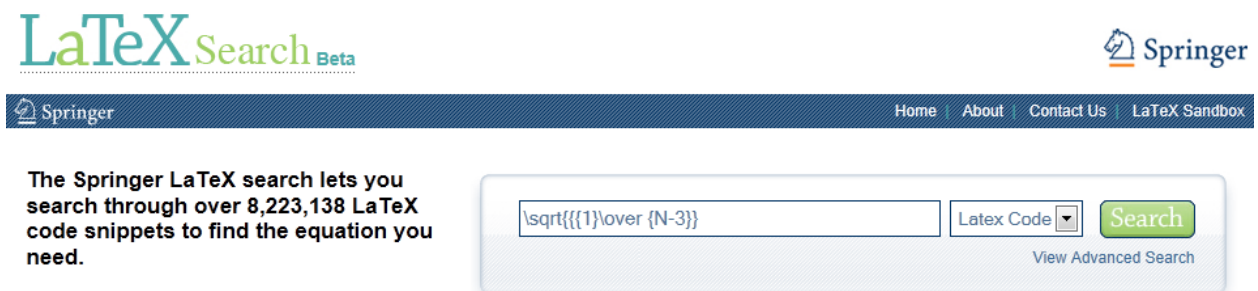
Figure 1: The Springer LaTeX search engine with a math expression entered in LaTeX.

Springer LaTeX Search was developed for researchers and allows the user to enter mathematical expressions as LaTeX code and submit their query to their database of scientific publications. Each result returned contains the LaTeX string searched for, a picture of the equation and additional information and links.

---

[1] http://www.latexsearch.com

Math Indexer and Searcher (MIaS) (Sojka & Liska, 2011)

Figure 2: The WebMIaS search engine with a math expression entered in TeX and the first 2 results shown.

WebMIaS is the web interface to the MIaS system and allows users to enter math expressions in TeX or presentation Math ML along with text. This math-aware search engine returns results that contain links to indexed documents along with match snippets to allow users to determine relevance. Similar to Springer LaTeX Search, it is designed to search scientific documents (users can choose from two different indexes) and will look for similar expressions that may be notated differently. Multiple expressions can be searched for at the same time.

EgoMath (Misutka & Galambos, 2011)
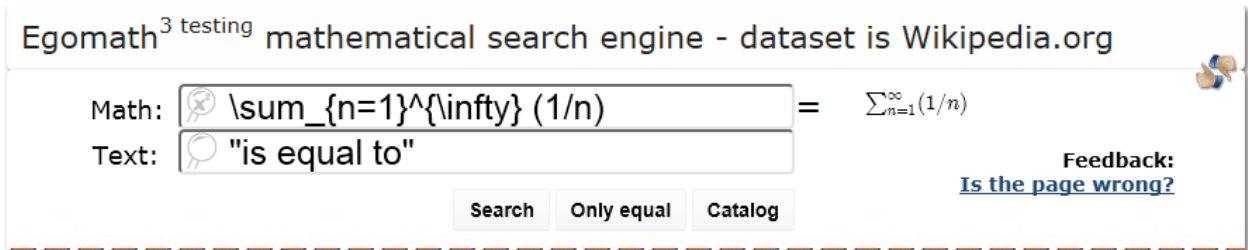
EgoMath is optimized for digital math content that is lacking in semantic information. Keywords have sufficient semantics for search engines to provide users with a relevant and sensibly ranked results list but, on their own, math symbols do not. Math symbols usually derive meaning from how they are organized structurally in a math expression (Kamali & Tompa, 2013). Because of EgoMath's focus on math content with little available semantics, its developers used it to enable math search in Wikipedia. Users can query using a math expression, text or both combined. Math expressions must be entered in TeX format in the "Math" search field and text is entered in a separate "Text" search field. TeX code entered into the "Math" search field is rendered into a formally notated math expression for the user's evaluation.

NIST Digital Library of Mathematical Functions (DLMF) (Miller & Youssef, 2003)

Figure 4: The DLMF search engine shown with a LaTeX equation entered and the first two search results.

The DLMF allows users to search with text and with mathematical expressions. Math expressions are entered as LaTeX or LaTeX-like code, using any of LaTeX's defined symbols and commands. It also allows the users to search metadata using descriptive text key words. An advanced search option allows users to restrict their search to specific chapters or restrict the type of results they want returned (e.g., equations, graphs, tables, etc.)

Wolfram Alpha[2]

---

[2] www.wolframalpha.com
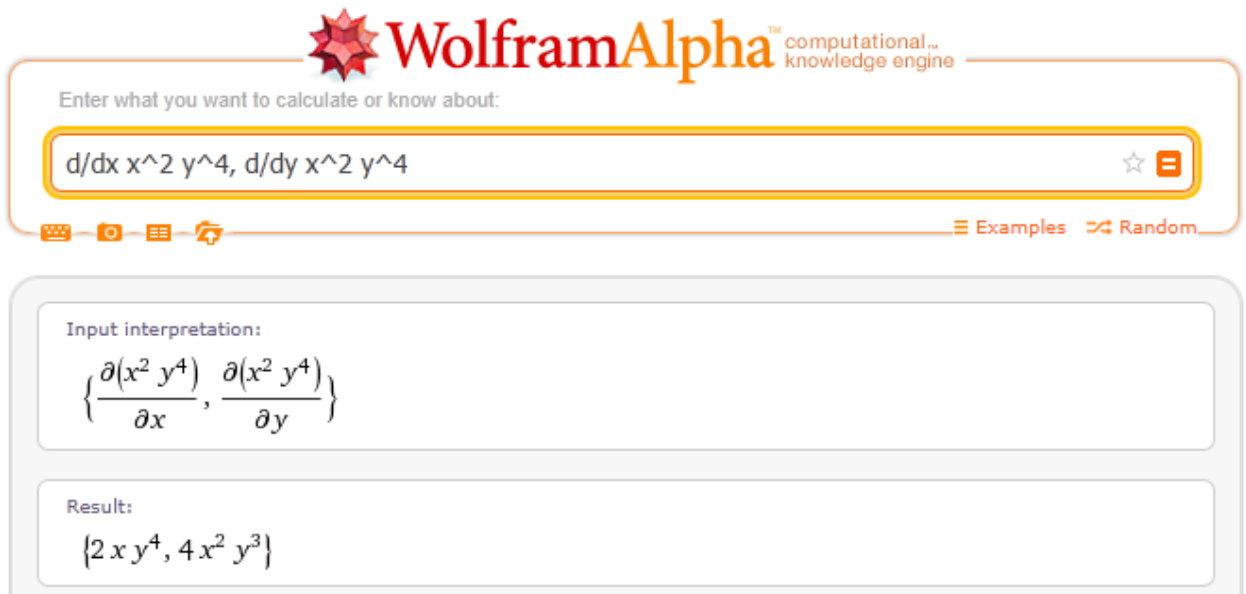
Figure 5: The Wolfram Alpha search interface showing a math expression entered in LaTeX format, along with Wolfram's interpretation of the formal notation and solution.

Wolfram Alpha is question-answering system designed to return answers to math problems or information from its knowledge base. It does not search the web. Users can input text queries or enter math expressions in LaTeX format as queries.

WikiMirs (Hu et al, 2013)

Like EgoMath, WikiMirs enables math search for Wikipedia. It is designed to look for textual and spatial similarities between the user's search query and the indexed information. The interface, which is currently unavailable online, accepts LaTeX input.

MathWebSearch (Kohlhase & Sucan, 2006)

MathWebSearch will accept math expression queries input as MathML or OpenMath format and is currently accessible with various front-end user interfaces, a few of which are shown below.

- **zbMATH[3]**

zbMATH
the first resource for mathematics

\int_?a^?b |?f(x)?g(x)| dx \leq ?r

$$\int_a^b |f(x)g(x)|\,dx \leq r$$
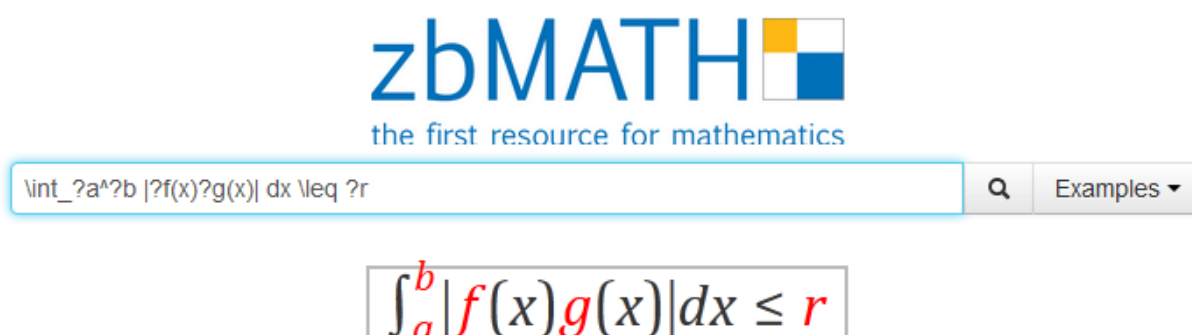
Figure 6: The zbMATH search interface with a math expression entered in LaTeX format and formal notation displayed beneath the search field.

This math search engine is powered by MathWebSearch and was developed to search the Zentralblatt Math database which consists of over 3 million entries from journals and books in the math and computer science domains. Users must enter their queries in LaTeX format.

- **Physikerwelt[4]**

---

[3] http://search.mathweb.org/zbl
[4] http://wiki.physikerwelt.de/wiki/Spezial:MathSearch
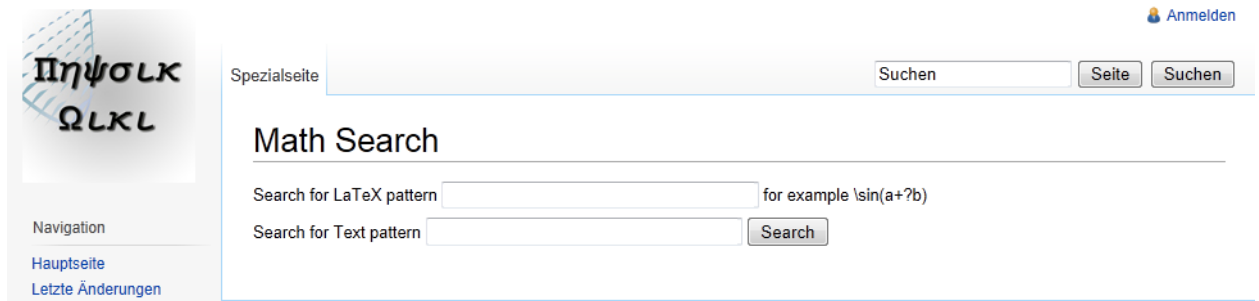
Figure 7: The Physikerwelt search interface.

This is another math search engine powered by MathWebSearch but its user interface is similar to EgoMath's in that it allows math expressions and text to be searched for simultaneously via input in two separate search fields. Math expressions must be entered as LaTeX.

- **SentidoSearch**[5]

---

[5] http://search.mathweb.org/sentido

Figure 8: The SentidoSearch math interface with an expression entered through use of the built-in equation editor.

This math search engine is also powered by MathWebSearch but it allows users to build math expression queries using a template editor.

MathFind (Munavalli & Miner, 2006)

Figure 9: The MathFind interface showing a math expression in the search field, built using the built-in equation editor.

MathFind searches math content by acting as an extension to a text search engine. The user interface allows math expressions and text keywords to be combined in a single search via two separate search fields. The user builds the math expression using the graphical equation editor provided in the interface and the expression is then converted to MathML for processing.

Math Aware Search Engine (MASE) (Nguyen, Chang, & Hui, 2012)
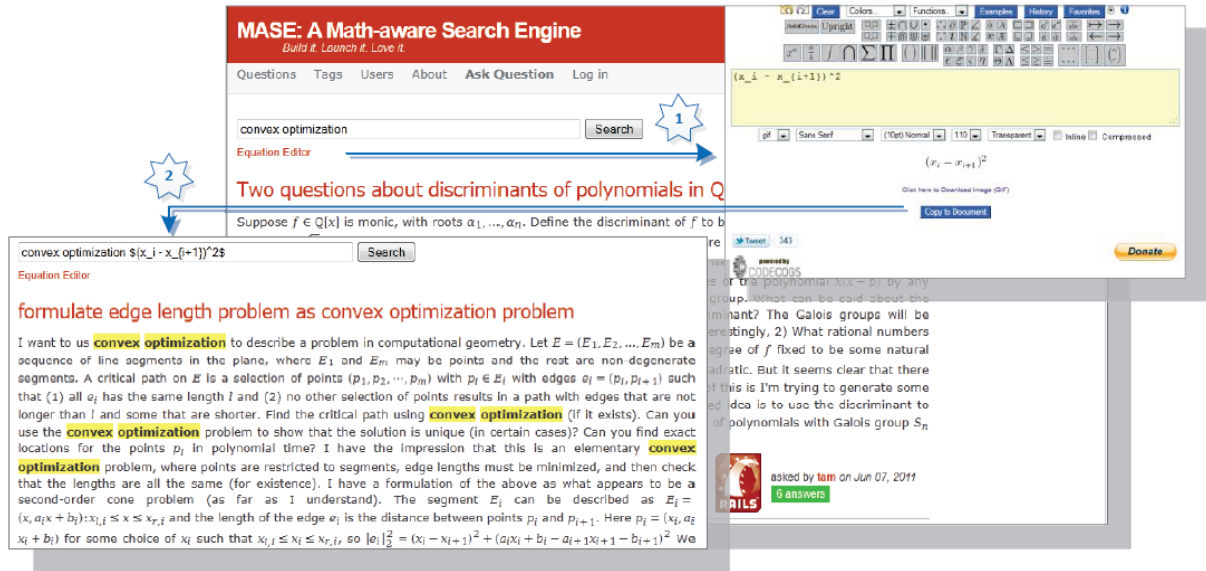
Figure 10: The MASE search engine with keywords and a math expression entered into the single search field, with the math expression built using the equation editor shown.

MASE is designed to support math question and answering (Q&A) systems. It allows users to input text queries and math expression queries using the equation editor CODECOGS as a front-end. The user builds the mathematical expression via the editor's interface and the editor converts the expression into LaTeX.

Symbolab Scientific Equation Search[6]

---

[6] http://symbolab.com/math/search
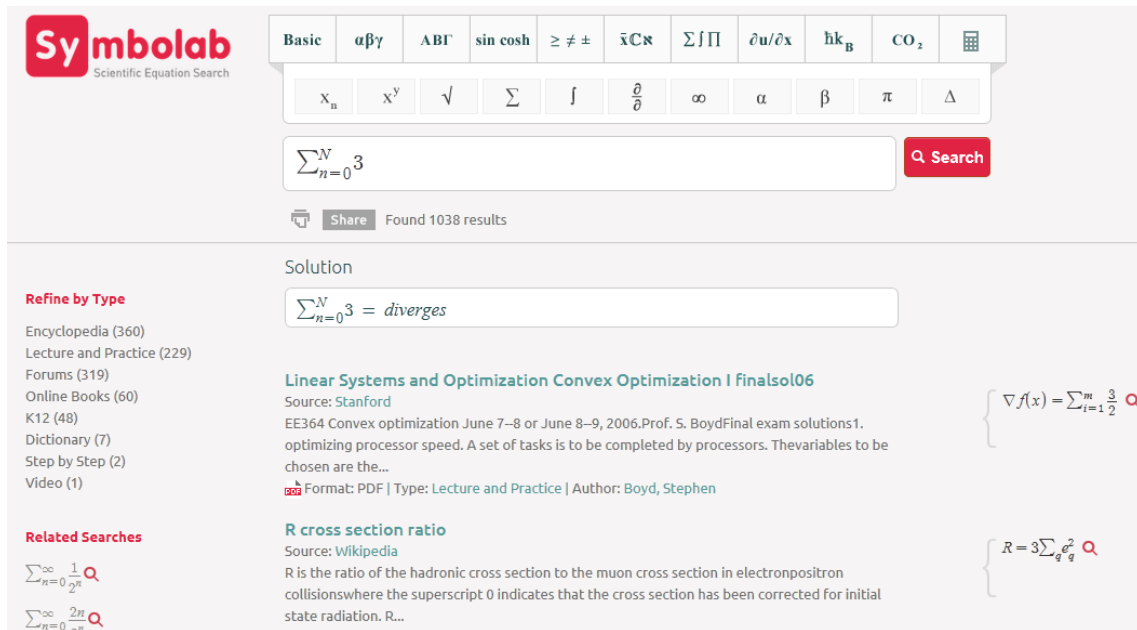
Figure 11: The Symbolab search engine shown with a math expression entered (using the equation editor), the first 2 results displayed beneath, and category refinements and related search suggestions displayed in the sidebar.

The Symbolab search engine, designed for math and science, interprets the meaning of the user's query by using proprietary machine learning algorithms. It not only searches but will also solve equations. The interface allows users to input text and math expressions in the same search field. The user builds math expressions using an editor. After search results are returned, the left sidebar allows the user to refine their results categorically and offers related math expressions that the user can search on by clicking an icon next to the related expression. Additionally, Symbolab offers a Quick Search plug-in for the Firefox and Chrome browsers (see image below). Once the plug-in is installed, whenever the user is on a supported site (Wikipedia, Wikibooks, StackExchange, or MathOverflow), all scientific content appears with the Symbolab icon next to it. Clicking the icon, submits the content as a search query on Symbolab's site.

Quadratic formula

A quadratic equation with real or complex coefficients has two solutions, called *roots*. may not be *real*.

Having

$$ax^2 + bx + c = 0,$$

the roots are given by the quadratic formula[1]

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

where the symbol "±" indicates that both

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad x = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

are solutions of the quadratic equation.[2]

Figure 12: Example of a Wikipedia page after the Symbolab Quick Search browser extension has been installed. Note the magnifying glass icons displayed next to the math expressions. Those can be clicked to submit the math expression to Symbolab's search engine.

2.3 Usability Concerns and Intention Gap

The math search interfaces presented above limit the user to expressing the math expression portion of their search need in the form of text, an encoding language, or via the use of some type of equation editor. Given the strong visual-spatial aspect of math notation and its impact on reasoning, and given the benefits of and preference for handwriting as a means of inputting math on a computer, it appears the current search interfaces may not provide the best user experience, particularly for non-experts, when searching for math-related information. If users are visualizing a formally notated math expression before they search and expect to be presented with results containing formally notated math expressions after they search, it seems unfortunate to have to break this visual flow by forcing the user to first convert the diagrammatic math expression into a coded, sentential one or build it through the tedious use of selection menus in order to initiate the search process. But that's exactly the behavior generated by the shortcomings of the current search interfaces.

Those who have studied user-centered design are familiar with Donald Norman's classic seven Action Steps that model users' behavior from goals and intent to action and evaluation.

1. Forming the goal

2. Forming the intention

3. Specifying an action

4. Executing the action

5. Perceiving the state of the world

6. Interpreting the state of the world

7. Evaluating the outcome

When there is a gap between a user's intentions and the system's allowable actions, it is known as the Gulf of Execution (Norman & Draper, 1986). This study hypothesizes that a Gulf of Execution occurs at the point where a user has formed the intent to search for information related to a math expression and yet the interface does not support a method for doing so in a natural way or in a way that matches the user's mental model. In the domain of search, the point at which a user is unable to precisely express their search intent as a keyword query has specifically been dubbed the Intention Gap (Zha et al, 2010).

As an example from another domain, music is a specialized domain in which information is not only learned, consumed and recalled in notation form, but also in auditory form. Yet, with the standard search interfaces, users interested in finding music with a particular tempo are limited to searching by text keyword. And while a trained musician may be able to input tempo information in text form, e.g., as a beats-per-minute (BPM) number, that kind of detailed information is not familiar to the ordinary user. This results in an Intention Gap which motivated researchers to explore multimodal input solutions to fill that gap and improve the user's experience by developing a search engine that allows users to play a sample piece of music, hum, tap, clap, etc. as a means of submitting a query (Yi, Zhou, & Wang, 2011).

2.4 m_{in} **interface** (Sasarak et al, 2012)

Acknowledging the shortcomings of existing math search interfaces, the m_{in} prototype was created to allow a more natural option for users to express their mathematical search need (Sasarak et al, 2012). m_{in} is a web-based interface that allows users to input a math equation by drawing it on a blackboard-like canvas using a mouse, pen or finger. m_{in} is multimodal – in addition to handwritten input, the canvas also accepts keyboard input and uploaded images.

Almost immediately upon receiving input, the m_{in} application displays its recognition of the user's input on the canvas, allowing for correction and/or manipulation. m_{in} offers expression manipulation tools such as symbol selection, stroke selection, and undo & redo, as well as an inline OCR correction menu.
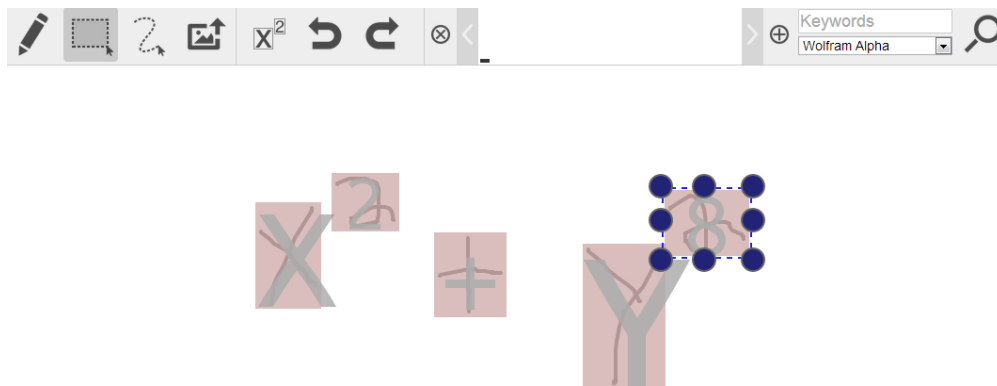
Figure 13: The m_{in} interface showing an expression drawn on the canvas and a symbol selected using the selection tool.
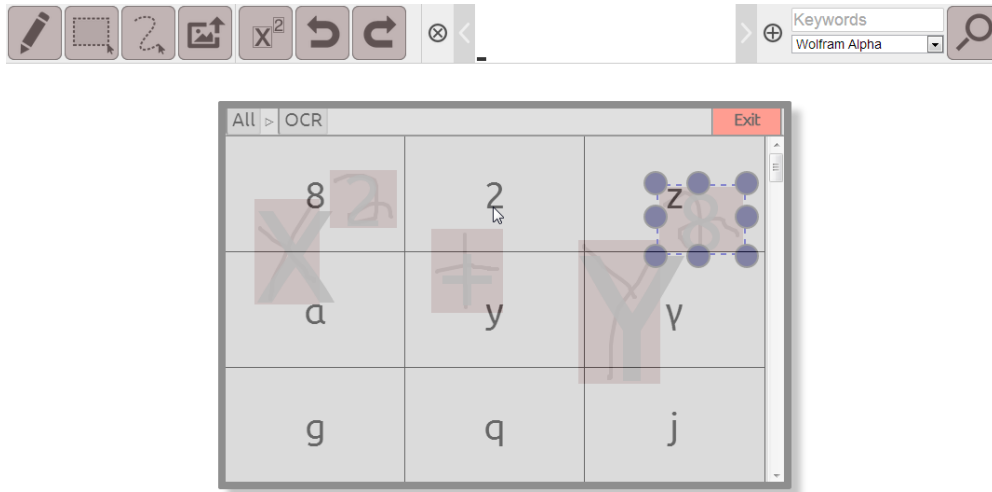
Figure 14: The OCR correction menu, accessed after double-clicking a selected symbol, allows the user a shortcut for substituting the selected symbol with another one.
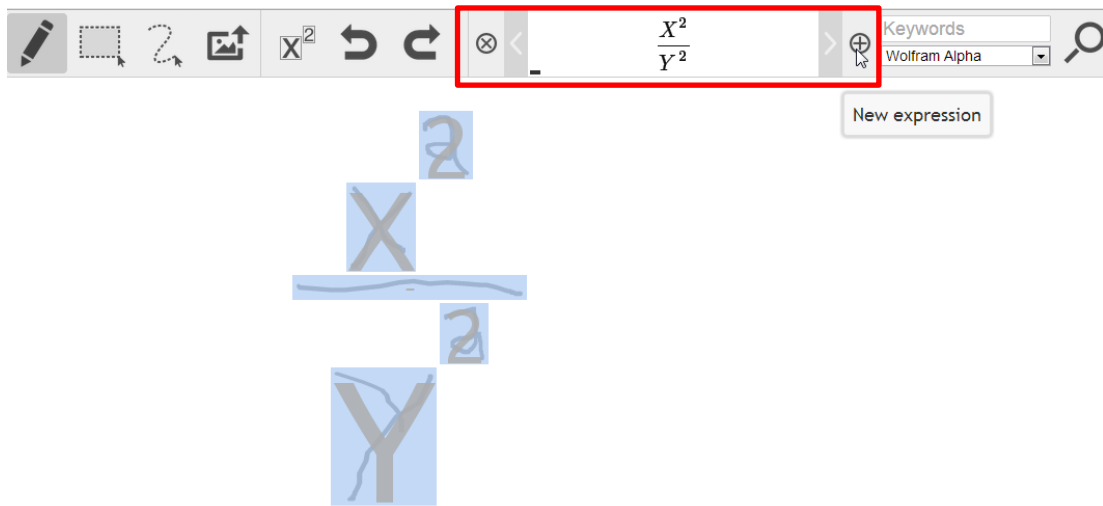
Figure 15: Shows an expression drawn on the canvas and then rendered in the slider area in formal notation after user has selected the Recognize Expression tool. Users may create a set of expressions by clicking the New Expression button.

Once the user is satisfied with the layout of the expression, m_{in} allows the user to submit the expression to various search engines directly from the interface. The interface also provides a conventional text search field allowing the user to submit keywords along with the expression.
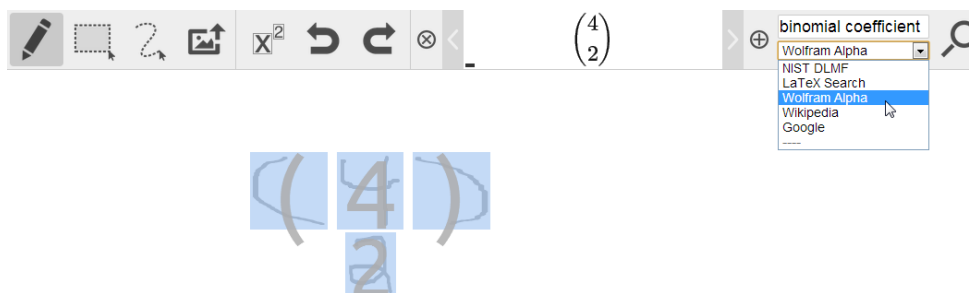
Figure 16: Shows a math expression drawn on the canvas and rendered in the slider area in formal notation, keywords entered in the text search field, and a selection of search engines the user can choose to submit the query to.

Scenario-based design

The research questions driving this study are exploratory in nature – an attempt to address such questions as "How do people use this new search interface?" and "How does this new search interface change search perception, motivation and behavior?" (Case, 2012). In keeping with the principles of scenario-based design, we wanted to "open ourselves to the real world" (Carroll, 2000) and learn by watching. So the study was observational, instead of experimental, in order to better understand the user's search behavior in the domain of mathematics.

The m_{in} search interface is novel and still in the prototype stage with some features still evolving. This makes it ideal for a technology-based scenario that will highlight any new affordances enabled by this novel interface (Carroll, 2000). So, in addition to observing how non-experts currently search for math-related information, the study is designed to take the additional step of actually putting multimodal query-by-expression technology in the hands of users to see what real-world scenarios can be discovered and to note any behavior changes introduced by the technology.

2.5 Summary

Part of the motivation behind one of the interfaces discussed above, the DLMF (Figure 5), was based on the understanding that math search engines must enable users to more naturally express their math query using mathematical notation. However, the focus was on allowing users to write their math query in the same way they would write it in a scientific document – in other words, in a LaTeX-like syntax. So it was suggested that a basic goal of any math search engine was to provide an "intuitive yet expressive math query language" (Youssef, 2006). Years later, when attempting to evaluate the success of the DLMF search engine, the DLMF's web logs showed a difference in the behavior of users searching for text vs. those searching for math expressions: not only was searching for math used less but when it was used, users needed to work more to find results, for example, looking past the first page of results, reformulating their query, and making multiple selections from the results. It could not be determined from the logs, however, if the behavior was caused by inadequate search results or the inherent nature of using a math expression as a query (Miller, 2013). In this thesis, it is hypothesized that query languages are not, in fact, natural for non-expert math searchers and intends to explore this idea of "natural" expression of a math search need a bit deeper by observing the search behavior of non-expert math searchers both with and without the use of the m_{in} interface. The intent is to observe any change in search behavior when users are presented with a math search interface that provides affordances for the preferred handwritten method of math input and the familiar, diagrammatic aspects of math notation and discover any other new affordances as well as real-world use cases.

3. Methodology

This chapter discusses the protocol and design of the study including the target audience, test environment, tools, and study tasks.

3.1. Target Participants

Our goal was to obtain a pool of participants consisting of non-expert math users. The ideal participant would fit the following profile

- 18 or older
- Currently enrolled in a first- or second-year college math course
- Self-rated as Beginner or Intermediate level in math knowledge
- Self-rated as Comfortable or Very Comfortable using the internet

For the purpose of the study, "Beginner" level math knowledge was defined as having knowledge of Basic Math and Pre-Algebra concepts such as arithmetic, powers and exponents, inequalities, polynomials and the Pythagorem Theorem.. "Intermediate" level math knowledge was defined as having knowledge of Geometry, Algebra and Pre-Calculus concepts, such as polygons, linear equations, exponential functions, factorials and the binomial theorem. Although prospective participants would be asked to self-rate their internet proficiency level, the screening questionnaire would also ask participants to indicate the internet activities they engage in more than three times a week. This would be done as a check, under the assumption that frequent and varied internet usage is a good indicator of higher comfort level.

Target participants were to be currently enrolled RIT students in the College of Science, which contains the math department, or the College of Computing and Information Sciences. Participants were recruited via an email (Appendix A) that briefly described the study and the $20 compensation and then invited interested students to go online to complete a screening survey. The first eighteen students who met the desired profile were contacted by email and scheduled into one of the available study sessions. Two of the selected students were used to pilot the study and the remaining sixteen participated in the actual study.

Although the ideal participant would be self-rated as Beginner or Intermediate level in math knowledge, seven students who self-rated as Advanced were selected to participate in the study because other information from their screening survey (Appendix B) indicated they may be Intermediate. For example, their current course enrollment or the number and type of math concepts they indicated familiarity with was consistent with other students who had self-rated as Beginner or Intermediate.

3.2. Environment and Tools

The study was run over four days, with each of the sixteen participants being scheduled for a one-hour session. To test and refine the protocol, two pilot sessions were conducted a week prior to the actual study. In all cases, a half-hour was allowed between sessions to allow for completion of notes and setting up for the next participant.

The study was conducted in RIT's usability lab in Golisano Hall. The lab consists of a testing room and observation room. The testing room was equipped with a desktop PC with a monitor, mouse and webcam, a speakerphone for communicating with the observation room and a one-way mirror on one wall. The PC in the testing room had the Firefox 20.0.1 browser installed, the latest version available, as well as TechSmith Morae Recorder 3.3. An iPad 2 with a stylus was also available for use in the testing room. The observation room was equipped with a desktop PC with TechSmith Observer 3.3 installed, a monitor, and mouse.

TechSmith Morae Recorder 3.3 was used to record the sessions. Online surveys created with Google Docs were used to collect screening, background, and post-study data (Appendices B, E, and G, respectively). A moderator in the test room with the participant took notes during each session recording which method participants used first to find math-related information (e.g., notes, textbooks, web site, or online search), which search engines participants used, what input methods were used (e.g., keyboard or mouse), length of participant's initial query (e.g., one term, two terms, etc.), format of participant's initial query (e.g., keyword, drawn expression, LaTeX or combination), the number of query reformulations and any other observations about the participant's search behavior, likes, dislikes, and areas of confusion, as well as any system bugs. For fourteen of the sixteen sessions, an observer in the observation room marked session

recordings with task times and notes on interesting participant comments, questions and actions; for the two moderator-only sessions, task times were marked on the recordings later.

3.3. Pilot Study

The two pilot participants were recruited in the same manner as regular test participants; however, pilot participants were not required to meet the math level requirement of Beginner/Intermediate nor the requirement of being currently enrolled in a first- or second-year math course. The two pilot participants both self-rated as Advanced in mathematical knowledge, one was a Freshman currently enrolled in Multivariable Calculus and the other was a Sophomore currently enrolled in Applied Statistics. The pilot tests were conducted in the Document Pattern and Recognition Lab with a moderator in the room and one person observing remotely via Adobe Connect. The sessions were recorded with Adobe Connect. The pilot participants were required to complete the same consent forms and complete the same pre- and post-study surveys as regular test participants.

Based on the pilot tests, the following changes were made to our protocol:

- Print each task on a separate sheet of paper so the participant can focus on one task at a time and so that it's easier for the moderator to rearrange the order of the tasks for the next participant.
- Have each participant read each task out loud so moderator can determine if the participant has misinterpreted the task before the participant starts working.
- At the beginning of each task, explain which tools can be used to complete the task.
- Be as modality-neutral as possible when providing assistance, being careful not to lead participants to draw an expression.

3.4. Protocol

Upon arrival for their scheduled session, participants were read an orientation script (Appendix C) and asked to sign a consent form (Appendix D) containing information about the study, the fact that the sessions would be recorded, how the data would be collected and used and who to contact about their rights as a participant. All participants consented to continue and were then directed to an online survey to complete a background questionnaire (Appendix E) designed to assess the participant's attitude toward math and to gather data about their current math study methods. This background data would later be used to help build profiles of the participants.

Participants were asked to bring any textbook, class materials or notes from their current math class. Most participants arrived with a backpack but it was not noted exactly how many, nor was it noted how many backpacks contained their math class text/materials.

The study tasks were designed to be consistent with math challenges non-expert students face in a freshman or sophomore college math course. Participants were reminded that they did not need to solve any math problems nor explain any of the math concepts to the study moderator – their job was to find resources that would help them explain the math concepts in each task to their classmate. Each task included both keywords and mathematical expressions to allow us to observe which the participant would use in their search.

Task 1

Your classmate is having difficulty recognizing polynomials. Find one or more resources to help explain to your classmate why $x^2 - 7x + 2$ is a polynomial and why $\frac{x^2-7x+2}{x+2}$ is not a polynomial.

Task 2

Your classmate has heard of Pascal's triangle but doesn't understand how it relates to math. Find one or more resources to help explain to your classmate how the equation $(x + y)^2 = x^2 + 2xy + y^2$ relates to Pascal's Triangle.

Task 3

Your classmate is struggling with binomial coefficients. Find one or more resources to help explain to your classmate how to find the value of ($\binom{4}{2}$).

Task 4

Your classmate is having trouble understanding the prime counting function. Find resources that help explain why $\pi(2) = 1$.

Participants were randomly put into groups of four. The tasks were counterbalanced across the groups to allow each task to be the first task for an equal number of participants and to present each task uniformly across the search tool conditions.

Group	Task Order
1	1 2 3 4
2	4 1 2 3
3	3 4 1 2
4	2 3 4 1

Search Tools and Conditions:

1. **Using their choice of text books, notes, websites, and/or online search.**

2. **Using online search only without the m_{in} interface.**

3. **Using the m_{in} interface only.**

4. **Using online search only with the option of using the m_{in} interface.**

In each session, between the second and third conditions, participants were shown the m_{in} interface, asked brief questions (Appendix F) to obtain their initial impressions, and then given a five-minute demonstration of m_{in}. The m_{in} interface supports handwritten input on the canvas (drawn with a mouse, pen or finger), keyboard input on the canvas, keyboard input into the search box and image upload. Of those, the input modalities available and demonstrated to participants included handwritten input on the canvas with a mouse, keyboard input on the canvas and keyboard input into the search box.

After each task, the following was asked of the participant and the response noted:

- How do you rank your success in finding one or more relevant resources?
    - 1 Unsuccessful
    - 2 Somewhat successful
    - 3 Successful

After completing all the tasks, participants were directed to an online survey (Appendix G) with Likert-scale questions allowing them to rate their experience. Upon completion of the survey, the participants were asked additional open-ended questions (Appendix H) by the moderator so participants could explain choices made or difficulties experienced during the session.

At the end of the session, participants were each paid $20 and provided with an information sheet about the study (Appendix I).

3.5. Summary

In summary, the study was designed to be observational rather than experimental. Our goal was to observe search behavior in non-expert math students, note whether that behavior was changed with use of a multimodal interface that allows both query-by-expression and keyword search, and discover any real-world use for such a tool. In essence, the goal was to learn by watching. The goal was not an experimental study to discover which search method was most effective or efficient. However, it was important to design the study and tasks rigorously to minimize anxiety in the participants, reduce bias in the results and ensure the correct things were being measured. To that end, the study was piloted, the tasks were counterbalanced, the sessions

were recorded, Likert-scale and open-ended questions were used and the tasks were designed with the help of a math professor. The study produced a large quantity of result data.

4. Results and Discussion

This chapter will present the more interesting and relevant results of the study, including the breakdown of participants, main findings regarding expression use, task time, task success, preferred input modality and real-world use for a multimodal interface that combines query-by-expression with conventional keyword search. This chapter also includes a discussion of the main results.

4.1. Participant Profile

The breakdown of actual participants was as follows:

Category	Attribute	Count
Gender	Female	11
	Male	5
College Year	Freshman	14
	Sophomore	2
Age Range	18-25	16
Math Level	Beginner	1
	Intermediate	8
	Advanced	7
Current Math Course	Data Analysis I	3
	Project-Based Calculus II	2
	Pre-Calculus for Engineering	2
	Algebra for Management Science	2
	College Algebra	2
	Calculus for Management Science	1
	Elementary Calculus II	1
	Probability	1
	Calculus B	1
	Engineering Statistics	1
Internet Proficiency	Very Comfortable	12
	Comfortable	4
Math Concepts Used in Study (how many participants have encountered it previously)	polynomials	16
	Pascal's Triangle	12
	binomial coefficient	11
	prime counting function	2

4.2. Expression Use in Search Query

There was a large increase in expressions used in the query when using m_{in} (see Figure 17). Before being introduced to m_{in}, only 2 out of 14 participants (14.3%) who chose online search in Search Condition 1 used an expression in their initial search query. None of the 16 participants used an expression in their initial query in Search Condition 2 where online search was required. However, when required to use m_{in} for the first time in Search Condition 3, expression use increased to 16 of the 16 participants (100%), with the majority experiencing at least one system or user error. 10 of the 12 participants who used m_{in} (83.3%) in Search Condition 4, where m_{in} was optional, used an expression in their initial search. None of the 4 participants who opted out of using m_{in} in Search Condition 4 used an expression in their initial query.
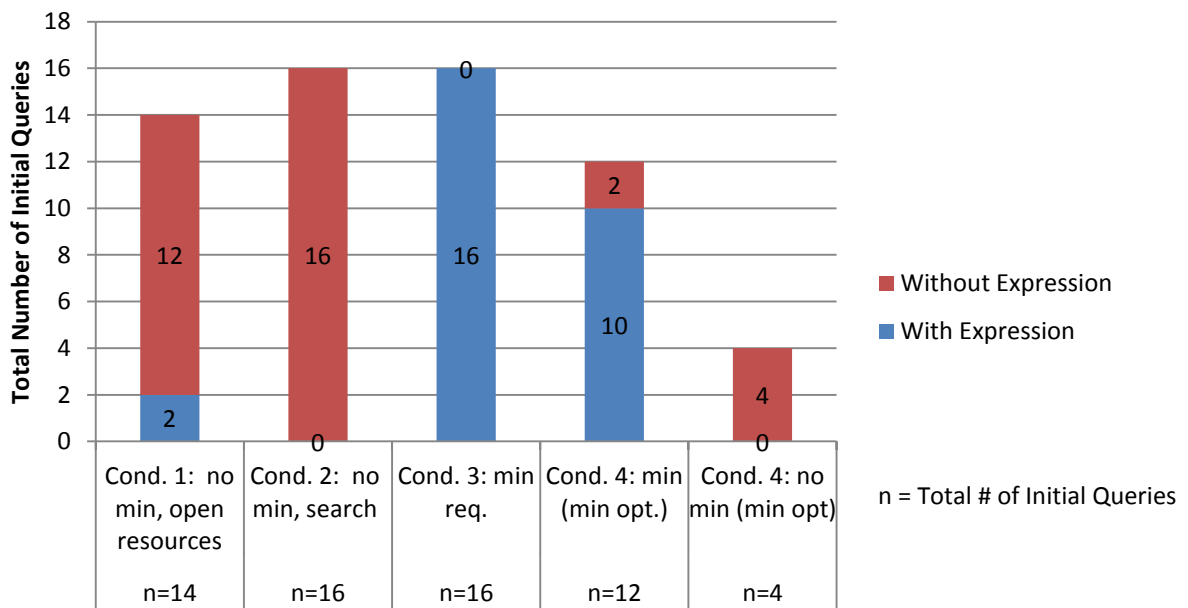
Figure 17: Initial Queries with an Expression vs Initial Queries without an Expression. Note that n=14 in Condition 1 because 2 participants did not do an online search, they used books & notes instead. Additionally note that data for participants using min and those not using min are displayed separately in the graph for Condition 4, where min use is optional, and have a combined n of 16.

Figure 18 shows that in non- m_{in} conditions (i.e., Search Conditions 1 and 2), participants used expressions in query reformulations more than they did in their in initial queries (Figure 17) and that a much greater percentage of the query reformulations contained expressions when

m_{in} was being used. An interesting observation regarding query reformulations is that when using *m_{in}*, participants would rarely remove the expression portion of their query once it was entered. The 2 behaviors that occurred more frequently were 1) participants would leave the expression portion of the query as originally entered and reformulate only the accompanying text keywords; and 2) participants would leave the entire query (expression and keywords) as originally entered and submit to multiple search engines. So, once an expression was entered in an initial query, it tended to remain in any reformulations the participant did.
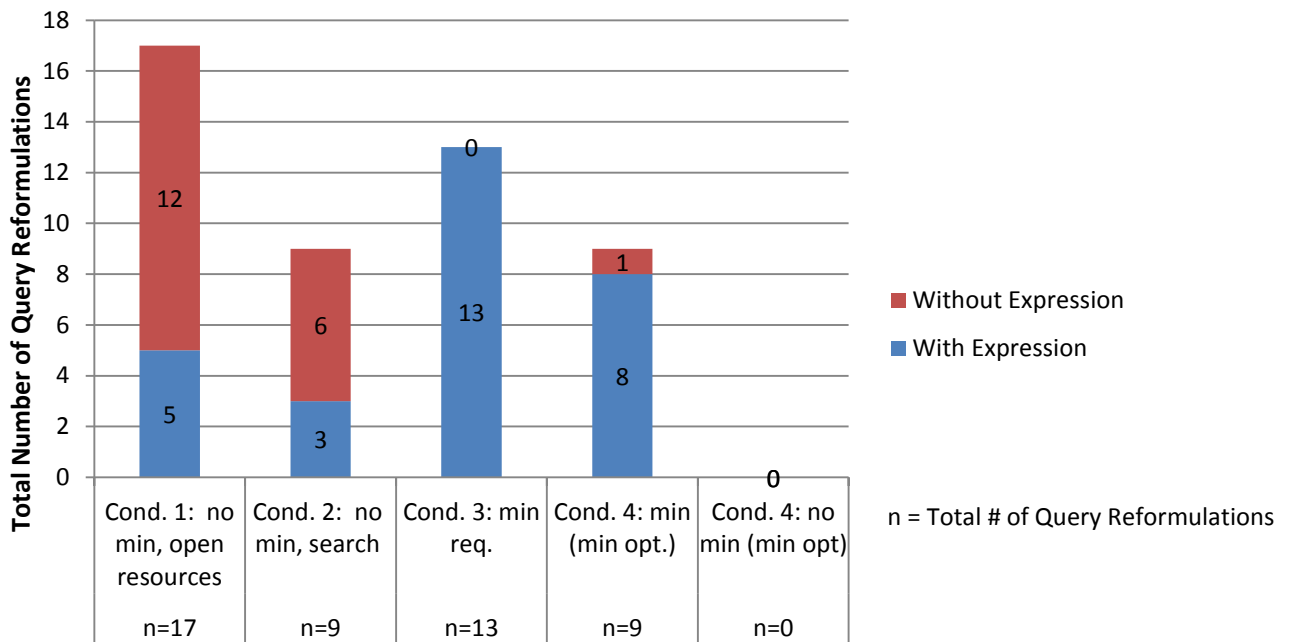
Figure 18: Query Reformulations with an Expression vs Query Reformulations without an Expression. Note that there is no max value for n as participants were allowed to reformulate as many times as they needed to. Additionally note that data for participants using min and those not using min are displayed separately in the graph for Condition 4, where min use is optional.

4.3. Task Time

Although task start and finish times were marked on the recordings during the session, the exact times were not noted until later during review of the recordings. For one participant, the system crashed twice and exact task times could not be recorded for two of that participant's tasks.

In general, there was a noticeable increase in average task time when participants used the m_{in} canvas. In Search Condition 1 using open resources, the average task time across all tasks was 147.4 seconds ($\sigma \pm 88.7$) and in Search Condition 2 using online search only, the average task time across all tasks was 118.87 seconds ($\sigma \pm 62.8$). Both of those non-m_{in} conditions averaged roughly between 2 to 2 ½ minutes. However, when required to use m_{in} for the first time in Search Condition 3, average task time across all tasks jumped to an average of 315.19 seconds ($\sigma \pm 239.8$) (Figure 18). This high average and variance in Search Condition 3 was caused by one participant spending almost 18 minutes on the Pascal's Triangle task (Task 2) and another spending 11 minutes on the Prime Counting Function task (Task 4), while participants completing the Binomial Coefficient task (Task 3) under the same condition were all able to finish relatively quickly in roughly 2 minutes or less.

When looking at average task time by task, rather than by condition, the times were almost always higher using m_{in} (Figure 19, green and purple bars). Worth noting, particularly for understanding how m_{in} is most useful and where it can be improved, is that the biggest increase in average task time was with the Pascal's Triangle task (Task 2). That task, whose expression contained several terms and had the most superscripts in the study, averaged 134.86 seconds ($\sigma \pm 74.25$) across both non-m_{in} conditions jumped to an average of 461 seconds ($\sigma \pm 273.28$) across both m_{in} conditions. In comparison, the Binomial Coefficient task (Task 3) had the smallest task time increase when using m_{in}. That task, whose expression ($\binom{4}{2}$) has a smaller number of terms and a relatively simple visual layout, averaged 121.57 seconds ($\sigma \pm 71.28$) across both non-m_{in} conditions and increased only 10% to an average of 134.29 ($\sigma \pm 45.05$) seconds across both m_{in} conditions. When completed under the m_{in} conditions, this task also had the lowest variance of all the per task time averages in the study – participants were more consistent in executing this task using m_{in}.
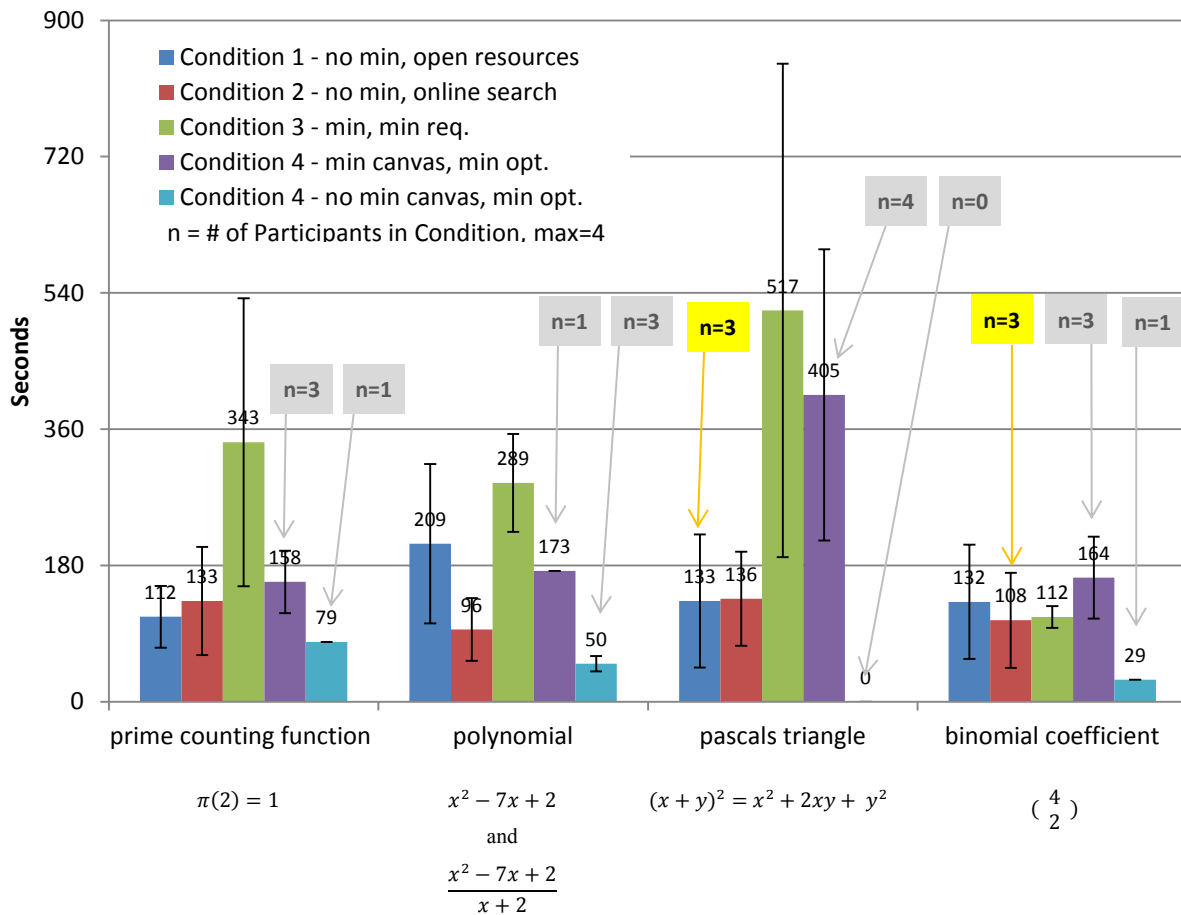
Figure 19: Task Times by Condition, Grouped by Task. n=4 unless otherwise noted. Note that data for participants using m_{in} and those not using m_{in} are displayed separately for Condition 4, where m_{in} use is optional, and have a combined n of 4. n is noted for each case (gray boxes). Due to recording issues, 2 task times could not be collected, resulting in an n of 3 for those averages (yellow boxes).

Although the expression ($\binom{4}{2}$) in the Binomial Coefficient task has a relatively simple visual layout, several participants had difficulty expressing it verbally and textually. In the post-study interview, a couple participants identified this expression as the one they were impressed at m_{in}'s ability to recognize correctly. Participant quote: "I was so surprised when it picked up on 4 choose 2." Another participant quote: "Like 4 choose 2 – that's really hard to 'write' but it knew what I meant and it accurately translated what I was trying to say to it."

4.4. Preferred Input Modality and Expression Format

Our results are consistent with the Anthony, Yang & Koedinger studies where participants found handwriting preferable to typing for inputting math on a computer. Upon first using m_{in}, 15 out of 16 participants (94%) used handwritten input on the m_{in} canvas (Figure 21). Additionally, in Search Condition 4 where using m_{in} was optional, 12 of the 16 participants (75%) chose m_{in} and 10 of the 12 participants (83%) that chose m_{in}, chose to use handwritten input on the canvas, despite the majority of participants having just experienced system or user errors and increased task time in Search Condition 3 (Figure 22). Participant quote when asked why handwritten input was chosen: "This is the way I'm used to thinking about math is writing it, it just seems more intuitive to write it."
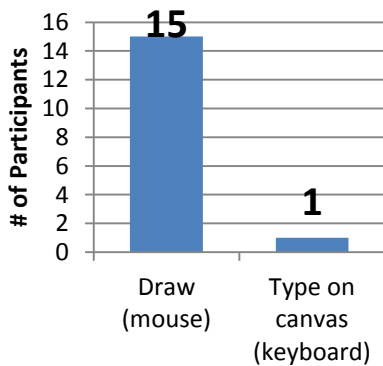

Figure 20: Input modality in Search Condition 3 where min is required. n=16.
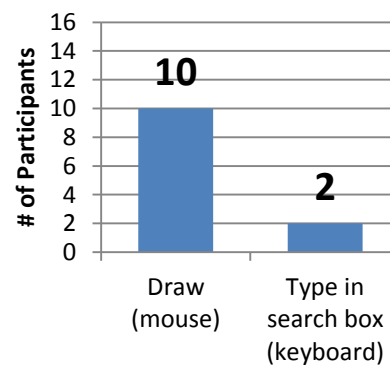
Figure 19: Input modality used in Search Condition 4 where min use was optional. n=12.

Our study observations confirmed the need to look at math experts and non-experts separately, and that math encoding languages and structure editors are not well within the grasp of non-expert math students. None of the participants were observed using any such methods for expressing their math search need. One participant made use of Wolfram's extended keyboard to enter a pi symbol and a few used standard math "shorthand" to indicate exponents (^) and division (/) or wrote out the word "pi" in place of the pi symbol. For the Binomial Coefficient task (4 choose 2), one participant attempted to mimic the visual layout and entered the expression into the search engine keyword box as the plain text query "(4 2)", was not satisfied with the search results, then opted for a keyword query instead. 11 out of 16 of participants (69%) agreed or strongly agreed that m_{in} made it easy to enter mathematical expressions, that m_{in} is a valuable tool, and that they would use m_{in} frequently when working on a math problem. 13 out of 16 (81%) agreed or strongly agreed that being able to search using mathematical expressions

made it easier to find what they needed. This is interesting, given that user self-ratings of search success were roughly the same.

4.5. Task Success

Participants were asked to rate themselves as Successful, Somewhat Successful, or Not Successful upon completion of each task. Interesting to note are that success ratings don't necessarily correlate to task time or difficulty. The task that took longest with m_{in}, Pascal's Triangle, shows the same level of success among participants when using m_{in} versus not using m_{in} (Figure 22). Additionally, with the Binomial Coefficient task, there were more "Somewhat Successful" ratings when using m_{in} as opposed to not using m_{in}, yet, in the post-study interviews, several participants specifically mentioned that expression as one they were impressed at m_{in}'s ability to recognize correctly.

Prime Counting Function

Polynomial

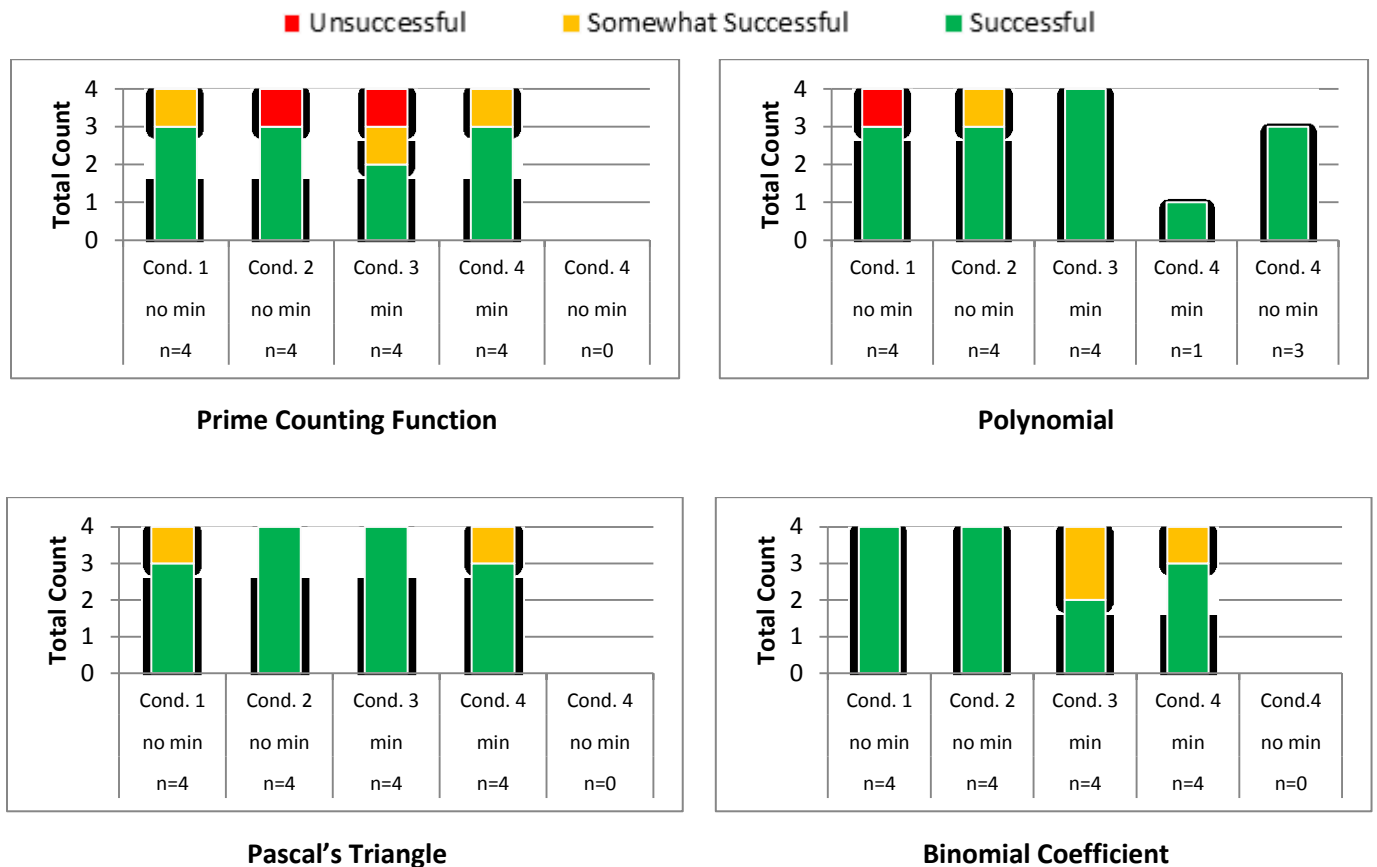**Pascal's Triangle**

Binomial Coefficient

Figure 22: Task Success by Task. n=number of participants. Note that data for participants using min and those not using min are displayed separately in the graph for Condition 4, where min use is optional, and have a combined n of 4.

4.6. Real-World Use

Participants were able to identify real-world scenarios where a search interface supporting query-by-expression would be beneficial for mathematical non-experts. In post-study interviews, 12 out of 16 participants (75%) identified scenarios where they would use m_{in} or could have used m_{in} in the past. This included studying for math tests (in particular, working with Calculus, integrals, complex math problems and expressions with lots of Greek letters), taking notes in class, collaborating with remote students on assignments, exporting expressions as image files or LaTeX for use in reports, and working with formulas from other domains such as accounting and biomedical research.

Most participants felt their ability to search improved using m_{in} and anticipated future benefits, even though average search task time increased using m_{in} and little difference was reported in task success rates. 69% of participants agreed or strongly agreed that the m_{in} application made it easy to enter mathematical expressions, that m_{in} is a valuable tool, and that they would use the m_{in} application frequently when working on a math problem. 81% of participants either agreed or strongly agreed that being able to search using mathematical expressions made it easier to find what they needed.

4.7. Summary of Results

The expression use and task time findings address the research question of whether the introduction of multimodal query-by-expression capability changes the search behavior of mathematical non-experts. The results of the study show that participants considerably increased their use of math expressions in their initial search queries and in their query reformulations when using the m_{in} interface and that task times increased with the introduction of m_{in}, particularly for the tasks involving expressions with more terms.

The findings regarding preferred input modality, user satisfaction and real-world usage address the research question of whether there are real-world scenarios where a multimodal interface that allows query-by-expression would prove beneficial for mathematical non-experts. The results of the study show there were benefits for the user for the metrics that were measured: user's satisfaction with using m_{in}, user's perceived value of query-by-expression and the m_{in}

interface, and user's anticipated usage of a tool like m_{in}. This study did not prove, nor did it seek to prove, that a multimodal interface is a faster way of searching or that it leads to more accurate search results.

In summary, we observed changes in search behavior when participants were allowed to draw expressions as part of their search query, and have identified a number of use cases for search interfaces supporting drawing expressions using a mouse/touch and keyboard. Additional results can be found in Appendices J – Q.

4.8. Discussion

We have presented a first study of math search behavior in math non-experts. This study was designed to target the non-expert math searcher as the feeling was that this is not a population familiar with LaTeX and other math encoding languages nor with template editors nor with expression names. Our observations may have proven that feeling correct – none of the participants were observed using any such methods (e.g., LaTeX, template editors) for expressing their math search need although a few claimed to be familiar with them. Admittedly, the structure of the expressions in 3 of the 4 tasks did not require any special encoding in order to be entered into a standard search engine text box. But for the one task that did, the Binomial Coefficient task containing the expression "4 choose 2", most participants not only did not attempt to use any special encoding languages, but many were not familiar with how to say the expression in words. They recognized it visually but beyond that, were unclear how to express it in sentential form – the only form afforded by the conventional search text box. This observation seemed to indicate that we had, indeed, targeted an audience where a tool such as m_{in} might be beneficial and also confirmed the need to look at math experts and non-experts separately and confirmed the suspicion that math encoding languages and structure editors are not well within the grasp of non-expert math students.

Clearly, the biggest behavior change noted in participants using the m_{in} interface was the drastic increase in the use of expressions in their search query. We feel that this change can be attributed to three things:

1. **The affordance of the interface.** The m_{in} interface is designed with a large open, blank area with several tools located horizontally across the top. Upon viewing the interface for

the first time, without using it or knowing its function, several participants mentioned how similar in appearance it was to a smartboard. When asked what the purpose of the interface was, the majority said "drawing." (Appendix M).

2. **The novelty of the interface.** When provided with a demo of the interface, several participants clearly showed visible and audible signs of being fascinated and impressed and seemed eager to try it themselves. In the post-study interview, about a quarter of the participants who chose drawing over typing cited "new and intriguing" as being their reason for making that choice.

3. **The ability for the interface to bridge the "intention gap".** Although participants may have chosen to draw their expressions on the m_{in} canvas because the open space was inviting and because the functionality was new and intriguing, the Binomial Coefficient task contained an expression that was difficult to enter correctly in a standard text search box. In the post-study interview, a few of the participants expressed not knowing how they would have entered it in a text search. Even though the "upload image" functionality of m_{in} was not demonstrated, several participants expressed really liking that feature, as well as suggesting real-world usage scenarios where they could imagine using that function of m_{in} for distance collaboration on math equations or sharing notes in class. All this feedback seems to indicate there are no interfaces readily available that allows this population to easily express math expressions in digital form.

As the m_{in} system is still in the prototyping phase, the majority of participants experienced glitches with the interface. Yet, the majority of participants rated their experience and the value of query-by-expression favorably. It seems the ease-of-use and performance of m_{in} only affected perceived value at the extreme end of the difficulty spectrum. In other words, a minor or moderate amount of difficulty did not detract from participants' perceived value but a high amount of difficulty resulted in negative ratings. This is probably due, again, to the novelty of the interface, but we also feel this is due to the interface's ability to bridge the intention gap better than any options currently available to them.

An interesting result is that with the Binomial Coefficient task, there were more "Somewhat Successful" ratings when using m_{in} as opposed to not using m_{in}, giving the impression that participants were less confident in their search results when using m_{in} for that task – the only task

where the expression contained within it could not easily be written in linear text form, a situation for which m_{in} is ideally suited. Ironically, it's the same task where, in the post-study interview, participants admitted to being very impressed that the m_{in} interface was able to correctly interpret the expression. So, what caused the lack of confidence in successful completion of this task? We believe the reason touches on something that might raise a concern about the impact of a tool like this on the non-expert population it is designed to assist. The reality is that the m_{in} interface correctly recognized what was written and passed along the correct information to the search engine but Wolfram didn't interpret the LaTeX code correctly. A division line was added, giving a result of 2 instead of the result shown in the task which was 6.
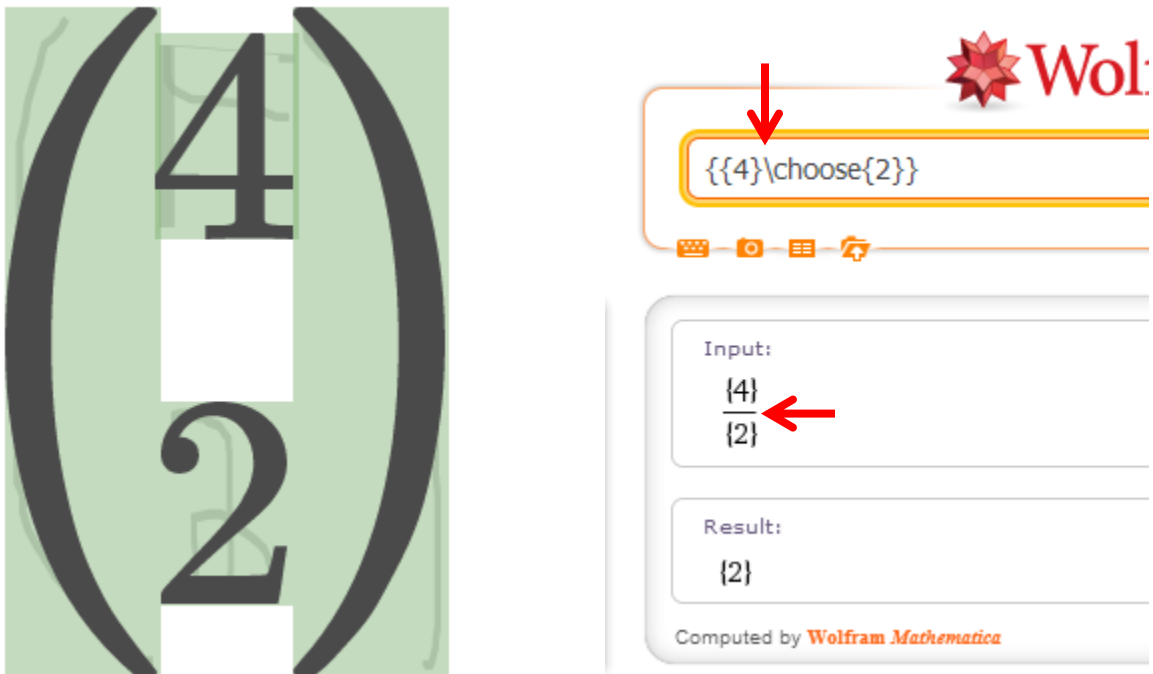
A couple of participants noted the discrepancy and removed the division line, thereby getting the same result shown in the task:

{{4}choose{2}}

Input:

$$\binom{\{4\}}{\{2\}}$$

Result:

{6}

Computed by **Wolfram** *Mathematica*
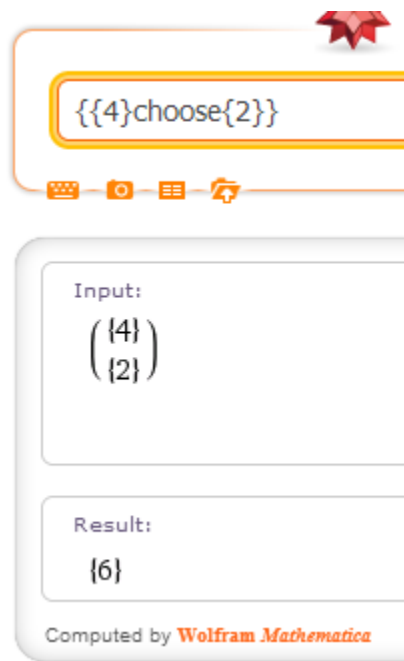
We feel that this discrepancy between what was input and accepted by the interface and what was returned in the search results caused an increased number of participants to feel only "Somewhat Successful" with their task completion, even though they didn't specifically point out the discrepancy. The fact that many participants didn't even notice the discrepancy highlights the over-reliance and trust a non-expert population might have with a tool such as m_{in}, believing that its innovativeness imparts a level of accuracy that may not be there.

Some interesting observations were made with regard to query reformulation. Of all participants who used an expression-only search, none seemed satisfied with only using an expression and all reformulated to add a keyword (Appendix J). This was an easy thing to do with m_{in} and may indicate positive benefits to an interface that allows the easy combination of query-by-expression and standard text keywords. Additionally noted, before the use of m_{in}, when participants reformulated a search query that contained an expression, they would make modifications to the expression, as well as any text keywords. However, when using m_{in}, when participants reformulated a search query that contained an expression, they left the expression as is and modified only the keywords. It could be that the difficulties encountered in inputting an expression made participants less likely to want to bother with the expression portion of the query again. But it could also be that because participants were able to express the expression in diagrammatic form – a form more compatible with their consumption of math expressions – that they were much more comfortable with the expression portion of the search query when using m_{in}.

There was also a behavior change noted in the way participants honed in on the resource they were looking for. Before m_{in}, participants would submit their search query and then "cherry pick" from the search results list until they found what they wanted. However, participants using m_{in} were more likely to submit their search to various search engines such as Google, Wolfram Alpha and Wikipedia. While submitting a search to various search engines is possible without m_{in}, that behavior was more likely to occur when using m_{in}. This behavior may have been enabled and encouraged by m_{in}'s design – the built-in drop-down list of search engines made it easy to submit the same search to various search engines. More research would need to be done to know whether this behavior led to better search results or not.

Since the goal of the study was to discover real-world usage scenarios of a multimodal interface and observe its impact on non-expert math searchers' behavior, the study was observational, not experimental. Task times were collected in order to provide additional perspective and/or insight into search behavior, not to test if one search method was faster than another. It was clear from the study that the using the m_{in} interface resulted in considerably longer average task completion times. Task time was most likely affected by bugs in the prototype and participant's unfamiliarity with its use. It was also clear that the m_{in} interface performed better with certain types of expressions than others. Longer expressions with numerous superscripts, such as those found in the Pascal's Triangle task and the Polynomial task, gave participants more trouble as there was more opportunity for them to encounter bugs in the prototype and have to start over. And those same expressions are pretty easily entered in sentential form using the standard characters on the keyboard. This is likely what led to some participants opting out of using m_{in} in Condition 4 where they had a choice.

The study informed changes to the m_{in} interface, including allowing operator shorthand in text input such as 'x^2,' which was typed by a number of participants. Typed and recognized expressions are now rendered using the online MathJax service (a TeX rendering engine) rather than our own algorithm which often produced unfamiliar symbol layouts. In the videos, participants frequently pause to understand the recognized expression, sometimes making comments about visual clutter causing the difficulty. As a result, handwritten strokes are now hidden after recognition and boxes are not drawn on top of recognized symbols (these visualizations are retained in edit mode). A method for correcting the grouping of strokes into

symbols was overlooked by some participants and not easily used by others – correcting OCR results required a press-and-hold gesture on a selected symbol and, along with interpreting the recognized expression, contributed to the longer task times when participants used m_{in}. That correction process has been replaced by a button for correcting symbol labels. The button brings up the correction menu. Based on study videos and questionnaires, improvements in the m_{in} interface have been made that may reduce observed increases in search task completion time when using m_{in} vs. text-based search. As most participants indicated query-by-expression through drawing to be valuable, search interfaces that support other diagram types (e.g., chemical diagrams) may be worth exploring.

5. Conclusion

The goals of this study were: 1) to observe whether m_{in} changes user search behavior, and 2) document relevant use cases for math search interfaces by mathematical non-experts. In conclusion, this study not only achieved those goals, but also makes several key contributions to the field of math search interface design. First of all, from the observations in this study, it seems non-experts in the math domain deserve special consideration in the design of math search interfaces. They are simply not familiar enough with math encoding languages and template editors to employ them readily in a math search query. Developers working on query-by-expression technology should feel encouraged to continue their work as it seems query-by-expression may be the missing piece that closes the intention gap in existing math search interfaces. Additionally, based on participant feedback and our observations, this study shows possible benefits in the implementation of query-by-expression capability alongside other input modes in a multimodal fashion. Researchers and developers in this domain should be careful to take note of the positive and negative ways in which a multimodal query-by-expression interface can change search behavior and then factor this knowledge into their interface designs. This study shows how search time, expression use, user confidence, user comfort, user satisfaction and query reformulation techniques are all impacted in interesting ways by the introduction of a multimodal query-by-expression tool that also allows conventional keyword search. Clearly, participants indicated such technology would be useful to them – developers can focus on decreasing negative impacts and highlighting positive aspects.

Future studies in this area should include testing the technology with expressions that have more complex structure and notation such as integrals, running experimental tests comparing query-by-expression with existing math search interfaces, exploring the effects of the technology with a younger population and populations from various educational and income backgrounds, longitudinal studies to understand usage patterns after the novelty of the technology has worn off, observing usage in the field as opposed to the lab, measuring actual search success (rather that user's perceived success), and exploring the use of this technology in other domains where search for diagrammatic terms and terms with non-standard text characters would be useful.

6. Bibliography

1. Adeel, M., Cheung, H.S., and Khiyal, S.H. (2005). Math Go! Prototype of a content-based mathematical formula search engine. *Journal of Theoretical and Applied Information Technology*, 1002-1012.

2. Anthony, L., Yang, J., and Koedinger, K. R. (2005). Evaluation of Multimodal Input for Entering Mathematical Equations on the Computer. *Proceedings of the SIGCHI Conf. on Human Factors in Computing Systems (CHI'05)*, 1184-1187.

3. Anthony, L., Yang, J. & Koedinger, K.R. (2007). Adapting handwriting recognition for applications in algebra learning. *Proceedings of the international workshop on Educational multimedia and multimedia education (Emme'07)*, 47-56.

4. Anthony, L., Yang, J., and Koedinger, K. R. (2012). A paradigm for handwriting-based intelligent tutors. *International Journal of Human-Computer Studies*, (70)11, 866-887.

5. Beyer, H., & Holtzblatt, K. (1998). *Contextual Design.* San Francisco, CA: Morgan Kaufmann Publishers, Inc.

6. Carroll, J. M. (2000). *Making Use: Scenario-Based Design of Human-Computer Interactions*. Cambridge, MA: The MIT Press.

7. Case, D. (2012). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior.* Bingley BD16 1WA, UK: Emerald Group Publishing Limited.

8. Hu, X., Gao, L., Lin, X. Tang, Z., Lin, X., & Baker, J. (2013). *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, 11-20.

9. Kamali, S., & Tompa, F. (2010). A new mathematics retrieval system. *Proceedings of the 19th ACM international conference on information and knowledge management*, 1413-1416.

10. Kamali, S., & Tompa, F. (2013). Retrieving documents with mathematical content. *Proceedings of the 36th international ACM SIGR conference on research and development in information retrieval*, 353-362.

11. Kohlhase, M., & Sucan, I. (2006). A search engine for mathematical formulae. *Proceedings of Artificial Intelligence and Symbolic Computation*, 4120, 241-253.

12. Kusano, K., Nakatani, M., & Ohno, T. (2013). Scenario-based interactive UI design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*, 391-394.

13. Landy, D., & Goldstone, R. (2007a). Formal notations are diagrams: Evidence from a production task. *Memory & Cognition (pre-2011)*, 35(8), 2033-2040.

14. Landy, D., & Goldstone, R. (2007b). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 720-733.

15. Miller, B., (2013). Three Years of DLMF: Web, math and Search. *Intelligent Computer Mathematics: Lecture Notes in Computer Science*, Volume 7961, 288-295.

16. Miller, B.R. & Youssef, A.S. (2003). Technical aspects of the digital library of mathematical functions. *Annals of Mathematics and Artificial Intelligence*, 38, 121-136.

17. Misutka, J., & Galambos, L. (2011). System Description: EgoMath2 as a Tool for Mathematical Searching on Wikipedia.org. *Intelligent Computer Mathematics: Lecture Notes in Computer Science*, Volume 6824, 307-309.

18. Muhammed, A. & Khiyal, M.S.H. (2011). Search Systems for Math Information Retrieval: A Survey. *Journal of Computing*, 3(3), 203-207.

19. Munavalli, R., & Miner, R. (2006). MathFind: a math-aware search engine. *Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval*, 735-735.

20. Nguyen, T., Chang, K., & Hui, S. C. (2012). A math-aware search engine for math question answering system. *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, 724-733.

21. Norman, D., & Draper. S. (1986). *User Centered System Design: New Perspectives on Human-computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

22. Sasarak, C., Hart, K., Pospesel, R., Stalnaker, D, Hu, L., LiVolsi, R., Zhu, S., & Zanibbi, R. (2012). m_{in}: A multimodal web Interface for math search. *Symposium on Human-Computer Interaction and Information Retrieval*, Cambridge, MA ([online](#)).

23. Smithies, S., Novins, K., & Arvo, J. (2001). Equation Entry and Editing via Handwriting and Gesture Recognition. *Behavior and Information Technology*, 20(1), 53-67.

24. Sojka, P., & Liska, M. (2011). The art of mathematics retrieval. *Proceedings of the 11th ACM symposium on document engineering*, 57-60.

25. Yi, Y., Zhou, Y., & Wang, Y. (2011). A tempo-sensitive music search engine with multi-modal inputs. *Proceedings of the 1st international ACM workshop on Music information retrieval with user-centered and multimodal strategies (MIRUM'11)*, 13-18.

26. Youssef, A. (2006). Roles of Math Search in Mathematics. *Mathematical Knowledge Management*, Volume 4108, 2-16.

27. Zanibbi, R. & Blostein, D. (2012). Recognition and Retrieval of Mathematical Expressions. *Int'l. Journal on Document Analysis and Recognition*, 15(4), 331-357.

28. Zha, Z.J., Yang, L., Mei, T., Wang, M., Wang, Z., Chua, T.S., & Hua, X.S. (2010). Visual query suggestion: Towards capturing user intent in internet image search. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 6(3).

29. Zhao, J., Kan, M., & Theng, Y. (2008). Math Information Retrieval: user requirements and prototype implementation. *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, 187-196.

7. Appendices

A. Recruitment Email

Subject: Seeking Participants for Math Search Study

The Document and Pattern Recognition Lab (DPRL) at RIT is looking for participants in a study to observe how math students search for math-related information. The findings will be used to help guide the development of tools that support the search needs of novice math students.

The study is expected to last 60 minutes. Participants will not be asked to solve any math problems and the tasks are not designed to assess math knowledge or skill. We will observe, note and record how participants carry out the tasks.

Participants will be paid $20 for their time.

If you would like to participate in the project, please go online and complete the participant screener. Link to screening survey:

> https://docs.google.com/a/g.rit.edu/forms/d/15kn2DQkDHfygQikXEEz5dCTb2Z2gfqh2Y1mWvQuN_6Q/viewform

We will contact you if you're selected. If you have any questions regarding the study please contact Keita Del Valle (kpd8205@rit.edu, (951) 963-4780). Any questions about your rights as a participant may be directed to Heather Foti (Associate Director, Human Subjects Research Office, RIT: hmfsrs@rit.edu (585) 475-7673) or Associate Prof. Richard Zanibbi (Principal Investigator, rlaz@cs.rit.edu, (585) 475-5023).

Sincerely,
Keita Del Valle
Graduate Student, M.S. HCI Program

Department of Computer Science, RIT
DPRL Web Page: http://www.cs.rit.edu/~dprl

B. Participant Screener

1. **What is your gender?**

   O Male   O Female   O Other _____

2. **What is your age range?**

   O Under 18   O 18 – 25   O 26 – 35   O 36 – 50   O 51+

3. **What is your current level of college enrollment?**

   O Freshman   O Sophomore   O Junior   O Senior   O Graduate   O
   Other_____

4. **Are you currently enrolled in a college math course at RIT?**

   O Yes   O No

   If yes:

   a. **Name of current math course(s)? (e.g., College Algebra, Pre-Calculus, etc.)**_____

   b. **Course number of current math course(s)? (e.g., 1016-200, etc.)**_____

5. **How would you rate your level of math knowledge?**

   O Beginner (e.g., Basic Math, Pre-Algebra)
   O Intermediate (e.g., Geometry, Algebra, Pre-Calculus)
   O Advanced (e.g., Calculus, Differential Equations)
   O Expert (e.g., Graduate-level math, Math Professor)

6. **Which math concepts have you encountered in your current or past curriculum? (check all that apply):**

☐ polynomials	☐ L'Hopital's rule	☐ slope intercept	☐ prime counting function
☐ Pascal's triangle	☐ differentiation	☐ binomial coefficients	☐ quadratic equation
☐ binomial coefficients	☐ Pythagorean theorem	☐ exponential functions	☐ scientific notation
☐ linear equations	☐ rational equations	☐ real numbers	☐ interval notation

7. **How would you rate your Internet proficiency level?**

○ Not very comfortable using the Internet
○ Comfortable using the Internet
○ Very comfortable using the Internet

8. **Indicate the Internet activities you engage in more than three times a week (check all that apply):**

☐ Social Networking   ☐ Gaming   ☐ Watching Videos
☐ Shopping   ☐ Sending/Reading Email   ☐ Banking/Bill Paying
☐ Searching for Information ☐ Reading (news, blogs)   ☐ Listening to Music

9. **Please indicate the devices you are most comfortable accessing the Internet with (check all that apply):**

☐ Windows Desktop   ☐ Mac Laptop   ☐ iPhone
☐ Windows Laptop   ☐ iPad   ☐ Android Phone
☐ Mac Desktop   ☐ Windows Tablet   ☐ Other: _____

Your Name: _____Contact Phone*:_____Contact Email*:_____

* Your contact information will *ONLY* be used to communicate with you regarding this study.

C. Orientation Script

Hi, my name is_____ and this is _____. We're both RIT students working in the Document & Pattern Recognition Lab. Thank you for participating in our study. First, I'd like to tell you about what we need you to do today. As I explain, please do not touch the computer or any of the materials before I indicate when you should start.

We are interested in how non-expert math students search for math-related information. During our session today, I will ask you to do several tasks and, for each task, will tell you what resources you may use. We will be taking notes as we observe you completing the tasks and, to assist us in more accurately documenting your actions, we will also be recording video and audio of you completing the tasks.

Please keep in mind that we are not testing you or assessing your math knowledge and you will not be not be given any math problems to solve. We are interested in observing your search behavior.

If you agree to participate in this study and be recorded, please read and sign this consent form which explains our anonymous data collection process, that your participation is voluntary with no penalty for refusing to participate or be recorded, and who to contact for more information about your rights as a participant.

[give consent form to participant to sign]

Thank you.

Before we begin, I need you to fill out a short pre-study questionnaire. Let me know when you are done.

[open background interview survey in browser]

Thank you.

Now, on these sheets of paper **[point to overturned packet of paper]** are the tasks you are to complete. You will do each task one at a time. Please don't look ahead at the other tasks or skip any of the tasks. Once I say "Please begin", read the first task out loud and begin working. When you feel you've completed the task, say aloud, "I'm done." Feel free to talk aloud as you complete the tasks. You may ask questions, but we may or may not answer them, as the goal is to observe the choices you make on your own when searching for math-related information.

Once you have completed the tasks, there will be a short post-study questionnaire for you to complete and a brief interview.

Do you have any questions before we begin?

Great. Now, in all of the tasks, we want you to imagine you are a tutor for a struggling classmate. Please turn over the packet and read aloud the first task.

D. Informed Consent

Informed Consent: Math search study
Department of Computer Science, Rochester Institute of Technology
Rochester, NY

This research project is concerned with developing improved methods for entering mathematical expressions into a computer, for the purposes of making it easier to enter, manipulate and search for mathematical notation.

The study is expected to take approximately 60 minutes to complete. You will be presented with a set of math-related search tasks to be completed using different approaches. You will then be asked questions about your experience performing the tasks. You will not be asked to solve any math problems and the tasks are not designed nor intended to assess your math knowledge or skill.

There are no significant risks associated with the study. We will observe how you carry out the tasks, take notes and make recordings. Data is being collected anonymously – data produced by your participation will not be associated with your name will only be available to the principal investigator and the graduate student who conducted the study. Study results will be summarized by the graduate student and may be published in journals and presented at conferences.

Your participation is voluntary. Refusal to participate or to be recorded will incur no penalty, and you may discontinue participation or decline to be recorded at any time without any loss of the participation stipend.

For questions regarding the research project, please contact Dr. Richard Zanibbi (Associate Professor, Department of Computer Science: rlaz@cs.rit.edu, (585) 475-5023). For questions regarding your rights as a participant in the project, please contact Heather Foti (Associate Director, Human Subjects Research Office, RIT: hmfsrs@rit.edu (585) 475-7673).

Participant Name: (Please Print) _____

_____
Signature

_____
Date

Participant ID:

E. Background Survey

Participant ID _____

Date_____Time_____

1. **What is your *primary* reason for taking college math?**

   O Degree requirement   O Degree elective   O For interest and/or enjoyment

2. **In general, how do you feel about math?**

   O Really like it   O Somewhat like it   O Neutral   O Somewhat dislike it   O Really dislike it

3. **How easy or difficult is math for you right now?**

   O Very difficult   O Somewhat difficult   O Neutral   O Somewhat easy   O Very easy

4. **Which of the following resources do you use to study for your math course(s)?** *(check all that apply)*

   ☐ Class Notes   ☐ Online Search   ☐ Flashcards   ☐ _____

   ☐ Practice Exams   ☐ Study Groups   ☐ Tutor   ☐ _____

   ☐ Study Guides   ☐ Assigned Readings   ☐ Textbook   ☐ _____

5. **How frequently do you need to look up mathematical information? (For example, definitions of functions, math symbols, plots, models, theorems or proofs)**

   O Rarely   O Once a year   O A few times a year   O Once a month   O Once a week   O Daily

6. **How frequently do you need to express mathematical notation when using a computer, such as for writing technical documents, doing online searches, or in using computer programs such as Matlab or Mathematica?**

   O Rarely   O Once a year   O A few times a year   O Once a month   O Once a week   O Daily

F. Pre-Demo Questionnaire

Participant ID _____ **Date**_____Time_____

Recorder Name _____

Before I explain and demo this for you, please take a minute to take a look at the interface. Without clicking on anything…

1. **What do you think the purpose of this interface is?**

2. **Does it look easy to use?**

3. **Have you seen any other interface or resource similar to this?**

4. **Do you think this will be useful to you?**

5. **Is there anything specific you'd expect to be able to do with this interface?**

G. Post-Study Questionnaire

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Being able to search using mathematical expressions made it easier to find what I needed.	○	○	○	○	○
2. **I would use the min application frequently when working on math problems.**	○	○	○	○	○
3. I found the m_{in} application easy to use.	○	○	○	○	○
4. **I would need support (help) while using the min application.**	○	○	○	○	○
5. **Various functions of the m_{in} application were well-integrated.**	○	○	○	○	○
6. **I would need to learn a lot about the m_{in} application before I could efficiently use it.**	○	○	○	○	○
7. **The m_{in} application is a valuable tool.**	○	○	○	○	○
8. **The look and feel of the m_{in} application is appealing.**	○	○	○	○	○
9. **The m_{in} application made it easy to enter mathematical expressions.**	○	○	○	○	○

H. Post-Study Interview

Participant ID _____ **Date**_____Time_____

Recorder Name _____

1. Based on their behavior during the tasks, ask users to comment on their experience with the following:

   - Formulating queries
   - Reformulating queries
   - Submitting queries
   - Evaluating search results

2. Based on their behavior during the tasks, ask participants to explain their choice/preference of:

   - search method (notes, text book, website, search engine, m_{in}, etc.)
   - search engine
   - query type (keyword, LaTeX, expression, combination, etc.)
   - input method

3. Ask participants to comment about being restricted to using online search, and specifically to the m_{in} application, in Tasks 3 and 4.

4. Specific to the m_{in} application, ask participants:

   - What were their expectations?
   - What did they like?
   - What didn't they like?
   - Was there anything they didn't understand?
   - What additional ideas do they have about the usage of m_{in}?
   - Are there any real-world situations where they would prefer to use an interface such as m_{in}?

5. Ask participant if they have any questions/comments regarding the m_{in} application, the study tasks or the study itself.

Thank participant for participating in the study and compensate him/her.

I. Project Information Sheet

Math Recognition and Retrieval Project
Document and Pattern Recognition Lab
Department of Computer Science, RIT

Thank you for participating in our study. The Math Recognition and Retrieval Project is an ongoing effort in the Document and Pattern Lab (DPRL) at RIT. If you are interested in learning more about the project, or later on finding papers produced as a result of this experiment, you may visit the DPRL web page:

http://www.cs.rit.edu/~dprl/

If you have any questions regarding the research project, please contact Dr. Richard Zanibbi (Associate Professor, Department of Computer Science: rlaz@cs.rit.edu, (585) 475-5023). For questions regarding your rights as a participant in the project, please contact Heather Foti (Associate Director, Human Subjects Research Office, RIT: hmfsrs@rit.edu (585) 475-7673).

## J. Use of Expression Only in Query (no keyword) – By condition

	Without m_{in}		With m_{in}	
	1st condition open resources	**2nd condition** online search only	**3rd condition** m_{in} only	**4th condition** m_{in} / online search
initial query	1	0	4*	4*
reformulation	1	0	0	0

All were reformulated to add a keyword

K. Use of Expression Only in Query (no keyword) – By Task

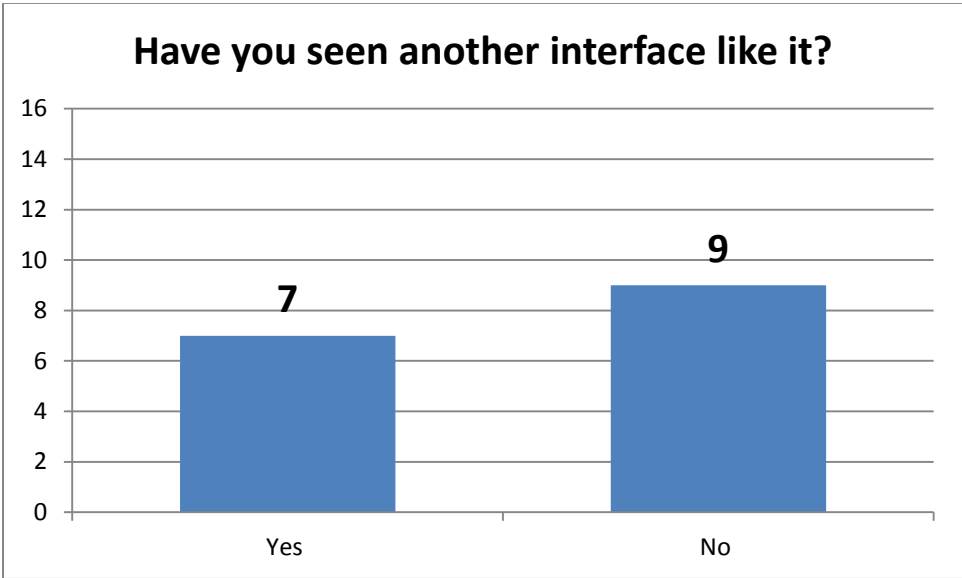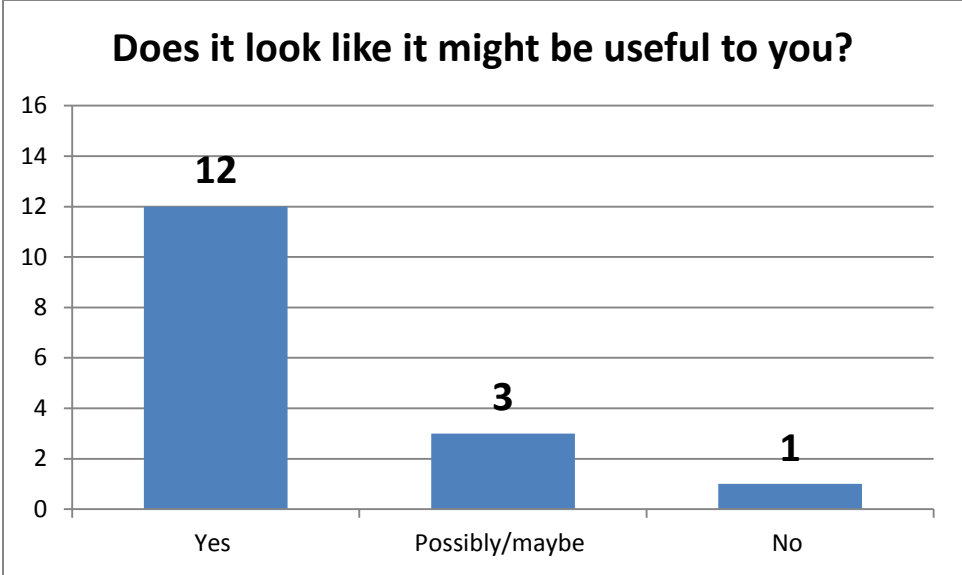|	Polynomial		Prime counting function		Binomial coefficient		Pascal's triangle	
	No m_{in}	m_{in}	No m_{in}	m_{in}	No m_{in}	m_{in}	No m_{in}	m_{in}
initial query	0	4	1	7	1	7	0	8
reformulation	1	6	5	3	0	5	1	4

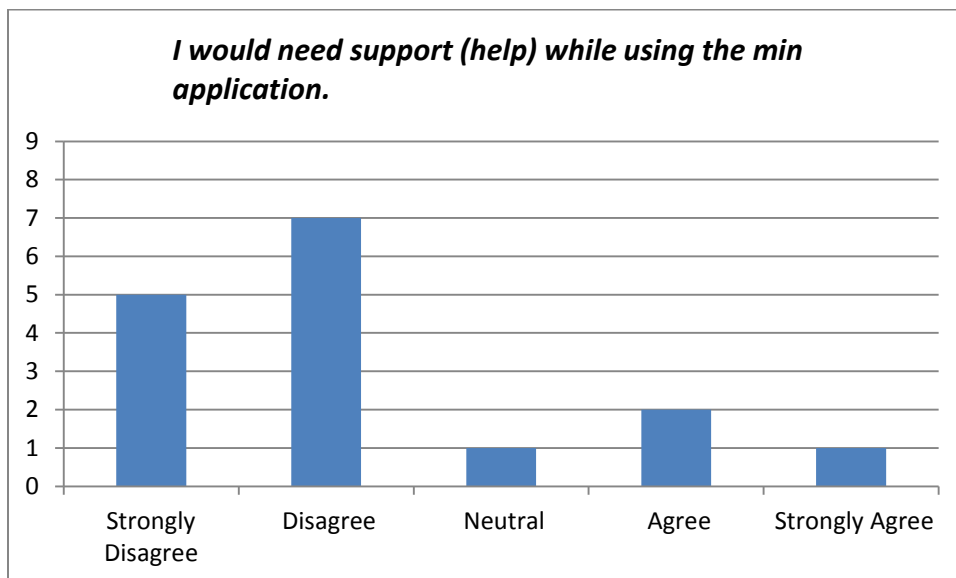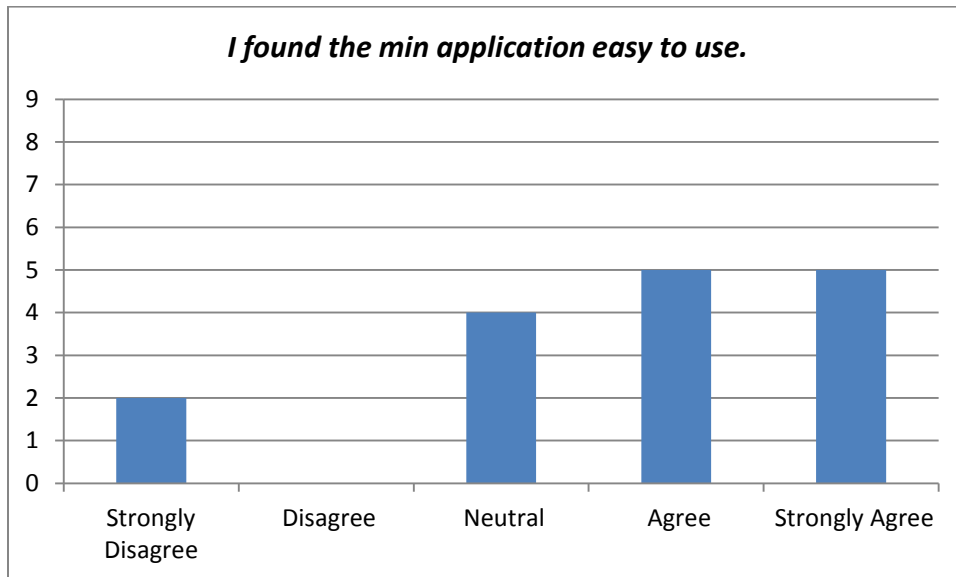## L. Participant Rationale for Choosing to Draw First
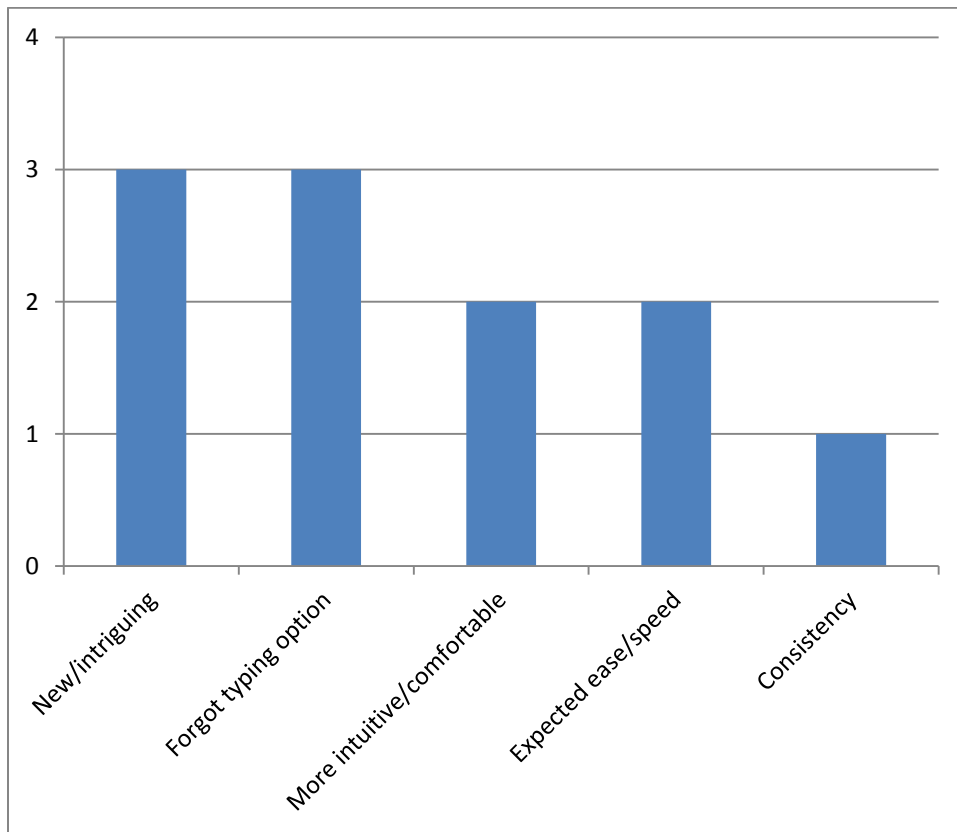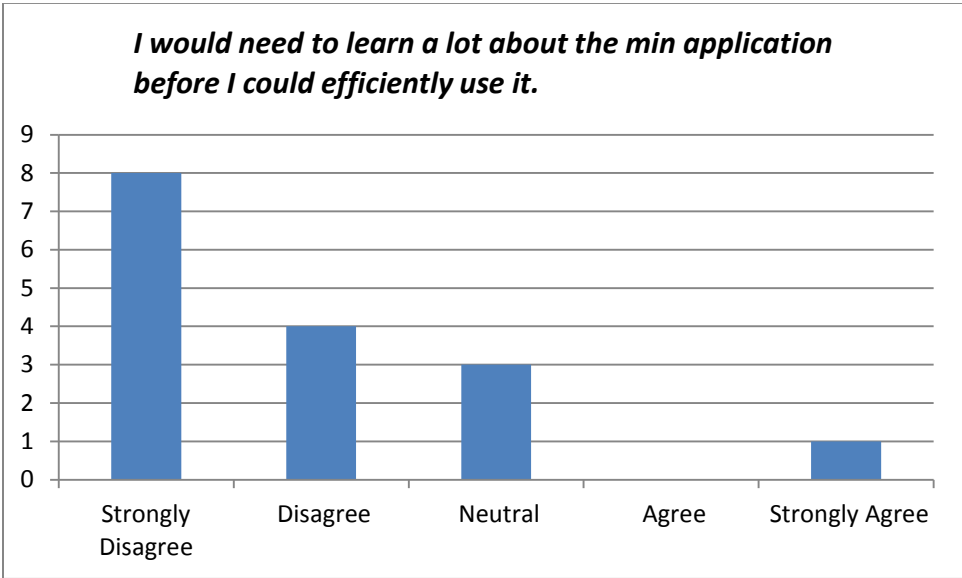
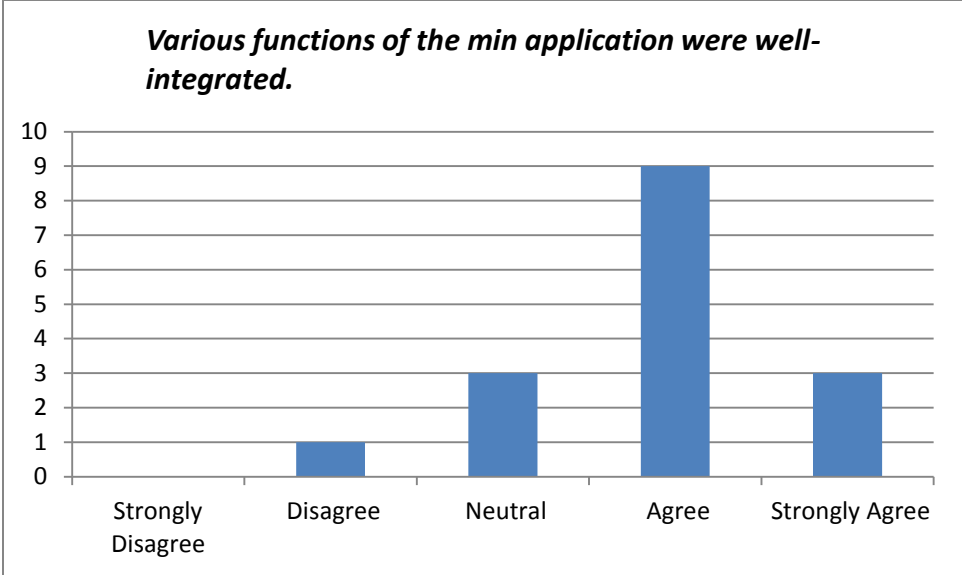## M. Summary of Participant Responses to Questions in Pre-Demo Interview

What do you think the purpose of the min interface is?

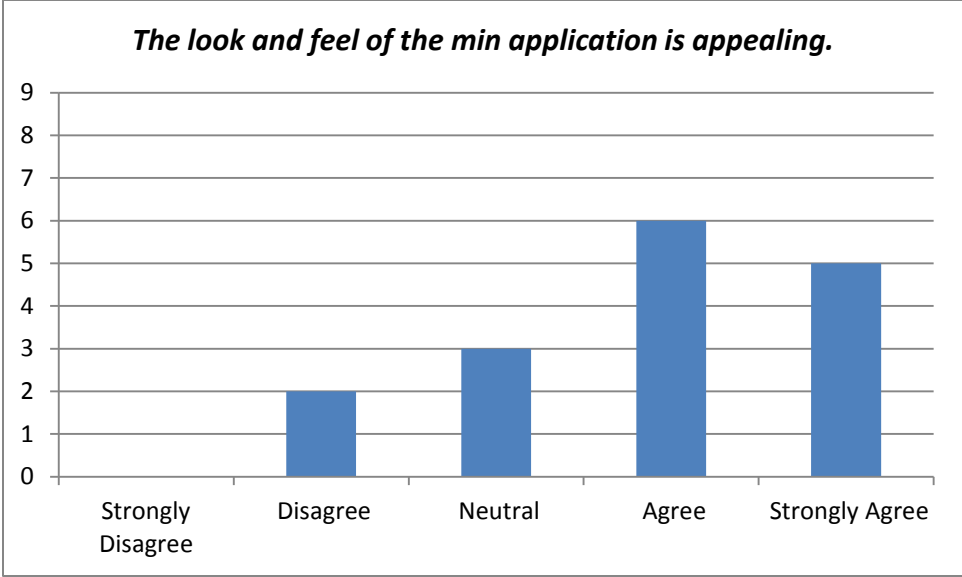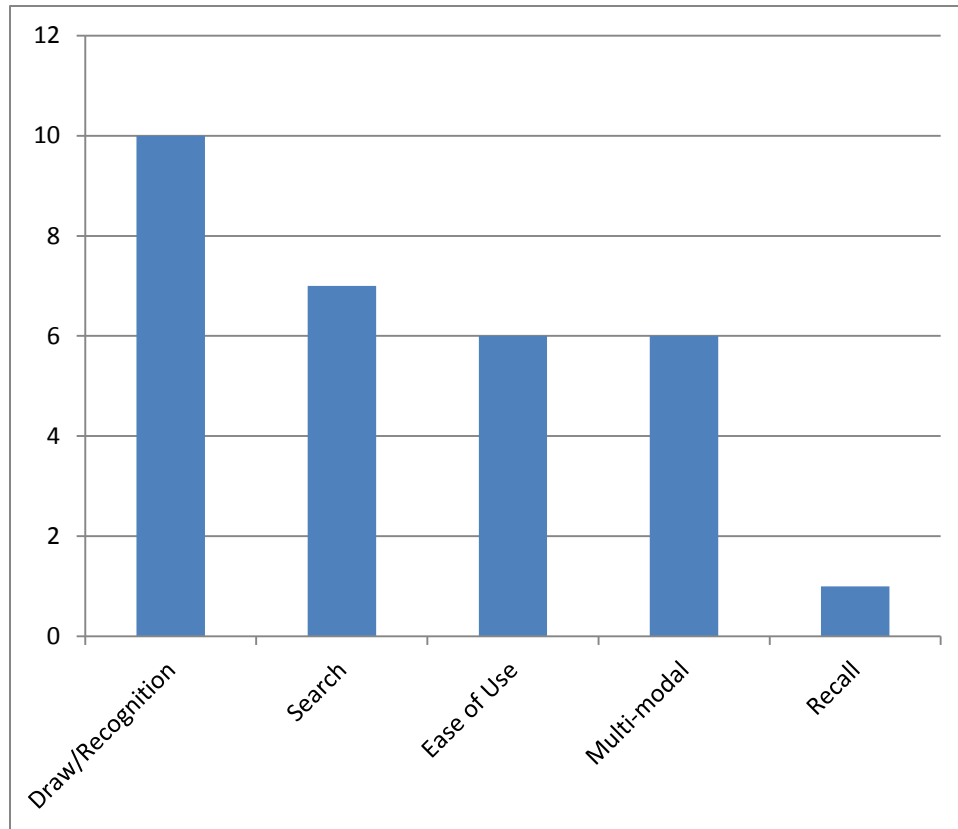## Does it look like it might be useful to you?

Response	Count
Yes	12
Possibly/maybe	3
No	1

Have you seen another interface like it?

Response	Count
Yes	7
No	9
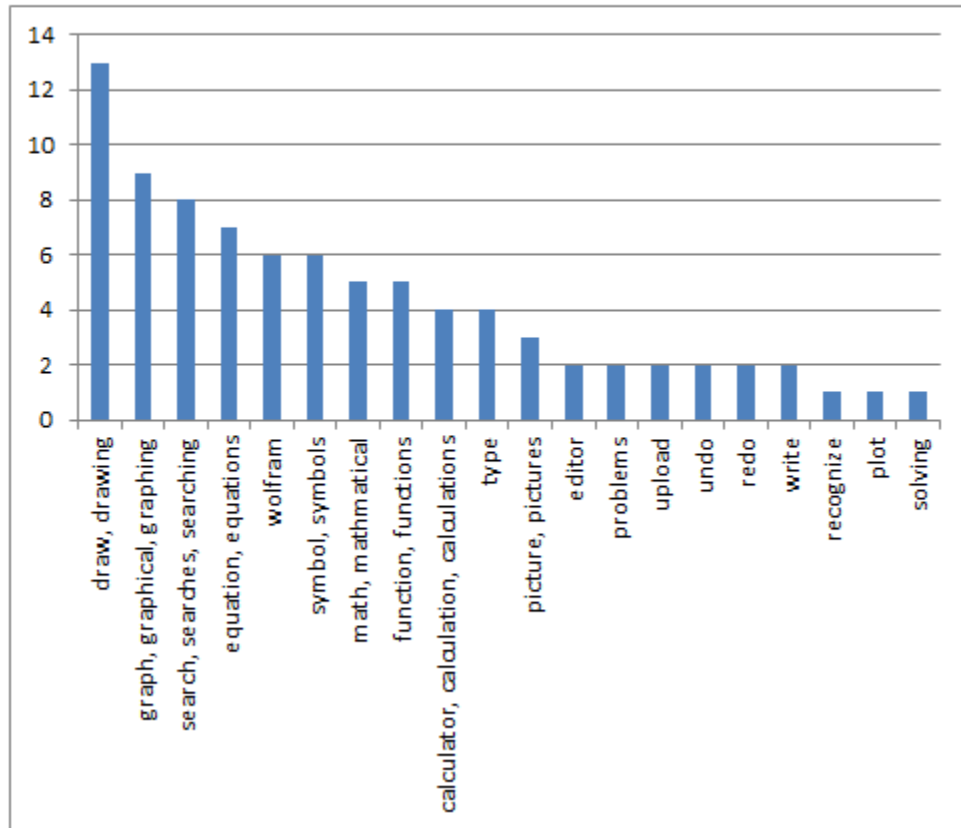
N. Summary of Participant Responses to Post-Study Likert Questionnaire

I found the min application easy to use.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Count	2	0	4	5	5

I would need support (help) while using the min application.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Count	5	7	1	2	1

Various functions of the min application were well-integrated.

I would need to learn a lot about the min application before I could efficiently use it.

The look and feel of the min application is appealing.
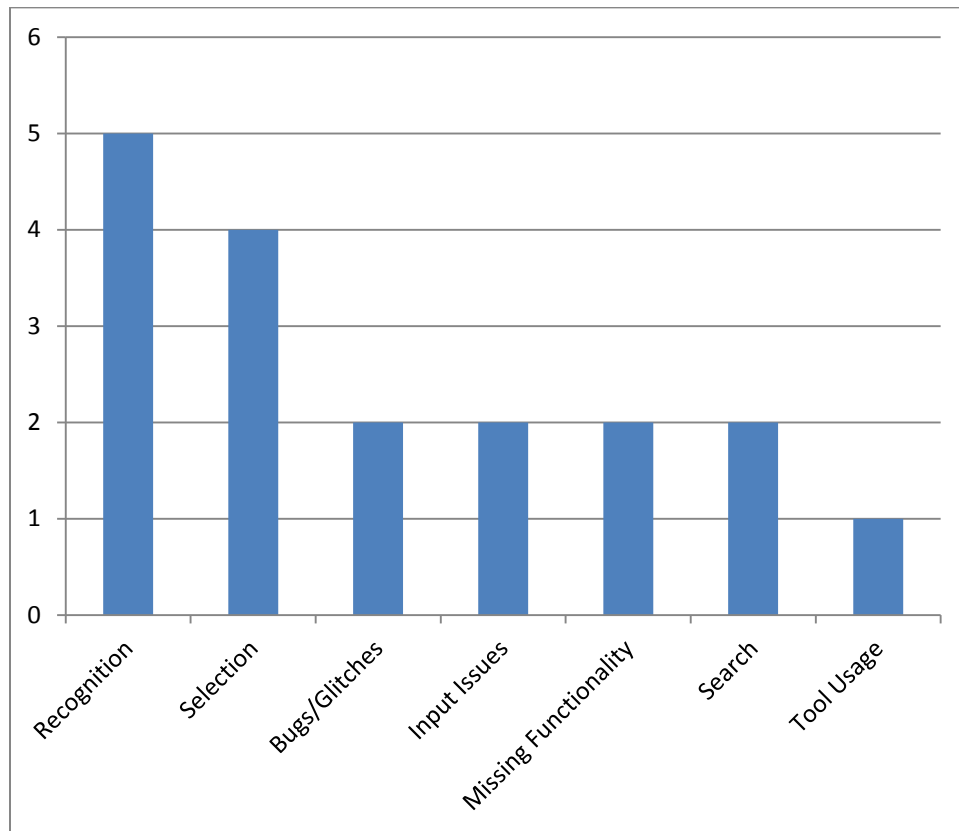
## O. Number of Likes about m_{in} by Category

- The 6 Ease-of-Use responses are further categorized as follows:
    - 4 - Easy
    - 2 - Efficient
- The 6 Multimodal responses are further categorized as follows:
    - 4 - Upload capability
    - 2 - Multiple input methods
- The 1 Recall response was in regard to the slider.

P. Number of Dislikes about m_{in} by Category

Q. Number of Suggestions for m_{in} by Category
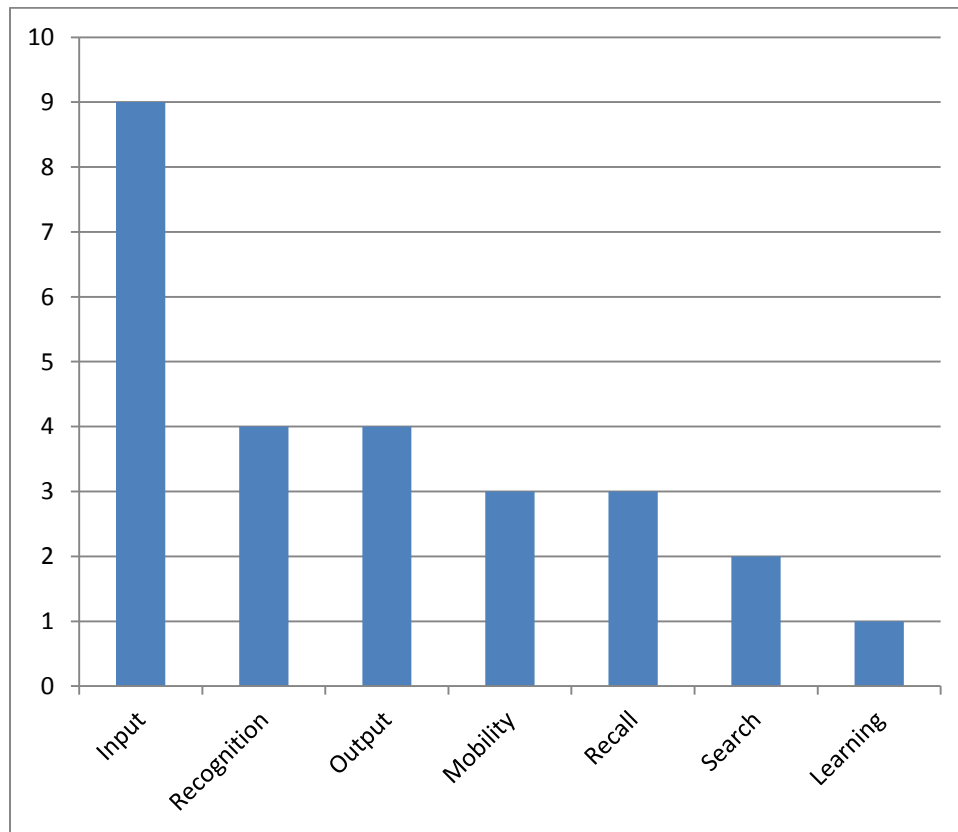
- The 9 Input responses can be broken down as follows:

    o 2 - Use of shorthand on canvas or slider

    o 2 - Type into slider

    o 2 - Voice recognition

    o 1 - Construction lines

    o 1 - Set scale of letters

    o 1 - Multiple colors

- The 4 Recognition/Rendering responses can be broken down as follows:

    o 2 - Other types of symbols/notation

- 2 - More Precision overlay/writing

- The 4 Output responses can be broken down as follows:

  - Exporting

  - Sharing/group work

- The 3 Mobility responses concerned the suggestion of providing a mobile app

- The 3 Recall responses can be broken down as follows:

  - 2 - Work/scratch area

  - 1 - Change the slider to drop-down

- The 2 Search responses concerned the suggestion of providing an option to add additional search engines to the available list

- The 1 Learning response concerned the suggestion of a Help section