CAMBRIDGE Mathematics

Glossary App

The development and pilot phase of *CM Define It*

Author

Dominika Majewska

Representing the work of

Lynn Fortin, Tabitha Gould, Rachael Horsman, Ellen Jameson, Vinay Kathotia, Ray Knight, Lynne McClure, Darren Macey, Dominika Majewska, Nicky Rushton, Lucy Rycroft-Smith and Ben Stevens.

CAMBRIDGE UNIVERSITY PRESS

UNIVERSITY OF CAMBRIDGE Faculty of Mathematics

UNIVERSITY OF CAMBRIDGE Faculty of Education

Cambridge Assessment

CAMBRIDGE
√Mathematics

Glossary App: The development and pilot phase of *CM Define It*

Sections in this document

CAMBRIDGE UNIVERSITY PRESS
UNIVERSITY OF CAMBRIDGE Faculty of Mathematics
UNIVERSITY OF CAMBRIDGE Faculty of Education
Cambridge Assessment

This report will describe a new development from the Cambridge Mathematics team – the *CM Define It* app – which is a research survey tool that collects information about existing definitions of mathematical key words. It will describe the aims, development and structure of the app as well as the pilot study which enabled the team to make adjustments to the app before its launch. The report will detail the changes that were made in response to participants' feedback post-pilot.

Introduction

Developing vocabulary is important to any experience that involves language (Monroe & Orne, 2002) and so knowing and understanding mathematical vocabulary is a key element of understanding mathematics (Miller, 1993, cited in Monroe & Orne, 2002). Mathematical vocabulary has been defined as "those words that label mathematical concepts (e.g. hexagon, dividend, and numerator)" (Monroe & Orne, 2002, p. 140). Riccomini, Smith, Hughes and Fries (2015) suggest that developing mathematical language is an essential element of teaching mathematics to young learners and this continues throughout an individual's journey in mathematics education. This is because understanding mathematical vocabulary enables the individual to access concepts and mathematical instruction (Monroe, 1998, as cited in Riccomini et al., 2015).

The availability and access to technology in education has steadily increased (Blackwell, Lauricella, Wartella, Robb, & Schomburg, 2013), affecting its usefulness and impact. In addition to the most commonly referred-to technologies (computers, laptops, tablets), internet-related technologies have transformed communication. For example, digital repositories enable creators to share their written products and help to create an interaction between authors and readers (Gómez, Cañadas, Soler, & Restrepo, 2009). Digital repositories are systems of software, hardware, data and procedures containing a range of documents, such as metadata and other digital objects, which allow for identification of the desired object through one identifier (Lopez, 2007, as cited in Gómez & Cañadas, 2013). They support the community by managing digital content and frequently offering open access, meaning that the online documents are free and do not have some of the copyright issues that other documents do (Gómez et al., 2009). One such example is *Funes* – a digital repository of documents in Mathematics Education (Gómez et al., 2009), which aims to improve the teaching and learning of maths in Latin America. Funes contains a taxonomy of key terms and aims is to support users in carrying out precise, relevant searches. It is also a flexible structure which evolves over time to meet the needs of the repository users; for example allowing users to assign labels freely to the documents, which in turn might encourage changes in

the taxonomy (Gómez & Cañadas, 2013, p. 8). This example of the use of technology is an important depiction of how technology can aid learning and research and how useful it is for a system to have key terms that can aid users in finding relevant material.

Some evidence suggests that wordlists are beneficial to both learners and teachers, as they support efficient teaching by providing teachers, learners, curriculum developers and test developers with a reference list of words that require attention. They may also help in deciding whether specific text is suitable for certain learners (Fahim, Fat'hi, & Nourzadeh, 2011). The highlighted benefits of wordlists support the aims of the glossary project described below and demonstrate the need for a holistic, accessible glossary of mathematical key words.

Over the last decade, the availability of educational apps has greatly increased (Douglas, Wojcik, & Thompson, 2011, as cited in Bouck et al., 2016). Apps have been designed for various educational uses including as a social-media style learning tool (e.g. Edmodo) (Bouck et al., 2016) and have therefore become a very popular method of collecting information in educational settings.

Aims of the glossary project

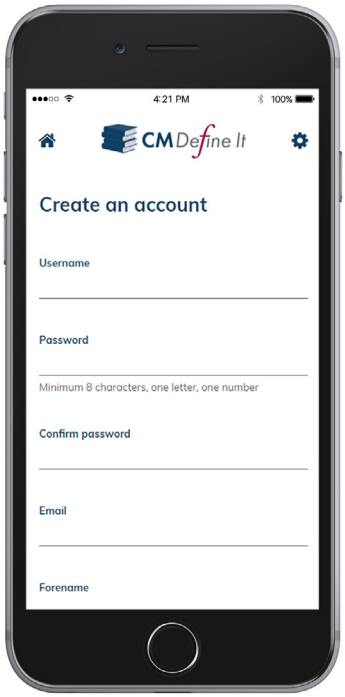
The *CM Define It* app is a survey tool developed to investigate how individuals in the mathematics education community perceive existing definitions of mathematical key words. The pilot study of the app examined what users would like to see and be able to do in an app that allowed them to rate mathematical definitions, what would make users engage with the app over time and what adjustments would make the app more user-friendly. The ultimate aim of the app is to collect information about what the mathematics education community perceives to be accurate, holistic and accessible definitions of mathematical key words. In the long-term, this will inform the glossary that will be attached to the Cambridge Mathematics Framework. For more information about the glossary layer in the Framework, please see the Ontology report (Jameson et al., 2019).

CM Define It

Development, structure and the pilot stage

The development of the *CM Define It* app began in November 2018. Initially we conducted research into the development of corpus-based dictionaries, language in mathematics, crowdsourcing and educational apps. A corpus refers to a "collection of written or spoken language data in a computer-readable format". It collates large quantities of language from many different real-life situations which are then used to compile dictionary entries (Oxford Dictionaries, n.d.). The research team also met and interviewed professionals working in fields such as dictionary development, classroom resource development and educational apps. This broad initial research narrowed down our focus and allowed us to start forming an idea of the literature which should support the app development, what the app should investigate and how it should look.

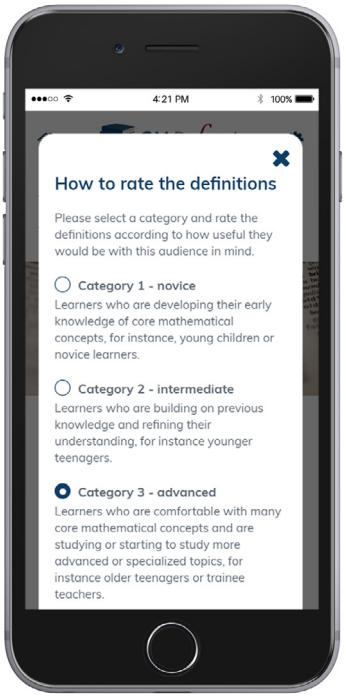
The *CM Define It* app contains a sign-up page. In order to sign up, users need to provide an e-mail address, some demographic information and their job role and must agree to our informed consent and privacy policy which adheres to the GDPR requirements.



Once logged in, users are presented with the word of the week – a mathematical key term selected by the Cambridge Mathematics team, which varies each week. We deemed it important to acknowledge that our audience may work with learners who are at different levels of learning and/or knowledge of mathematics. For instance, some participants may work with primary school students who have basic knowledge and/or understanding of mathematics. They may therefore be categorised as *novice* learners. In contrast, academic lecturers who teach undergraduate or postgraduate students may use more complex definitions of key words. They may therefore work with Category 3 (*advanced*) learners and thus prefer more detailed, complex definitions.

Before rating the definitions of the presented word, users were therefore asked to select a group of learners with whom they were most likely to use mathematical definitions. The three groups of learners defined for the purposes of this research were:

- *Novice learners* – learners who are developing their early knowledge of core mathematical concepts, e.g. young children,
- *Intermediate learners* – learners who are building on previous knowledge and refining their understanding, e.g. younger teenagers, and
- *Advanced learners* – learners who are comfortable with many core mathematical concepts and are studying or starting to study more advanced or specialised topics, e.g. older teenagers or trainee teachers



Haas, Pattuelli and Brown (2003) reported on the development of the Statistical Interactive Glossary (SIG) – a glossary of concepts in the field of statistics. The ultimate aim of the glossary was to increase users' understanding of statistical concepts and words. They highlighted that glossaries should address understanding of key concepts and that the audience should be "everyday users" (p. 194) rather than experts. They also suggested that glossary users should choose the style of explanation they find most effective. In addition, Ball and Bass (2002) claim that a mathematical definition is not useful if the terms do not line up with the users' knowledge, and that definitions should be mathematically appropriate and useful to learners who are at different levels. This demonstrates the importance of distinguishing between different categories of learners, giving users the opportunity to choose the category they work with and ensuring that the range of definitions in the app addresses these different groups of learners.

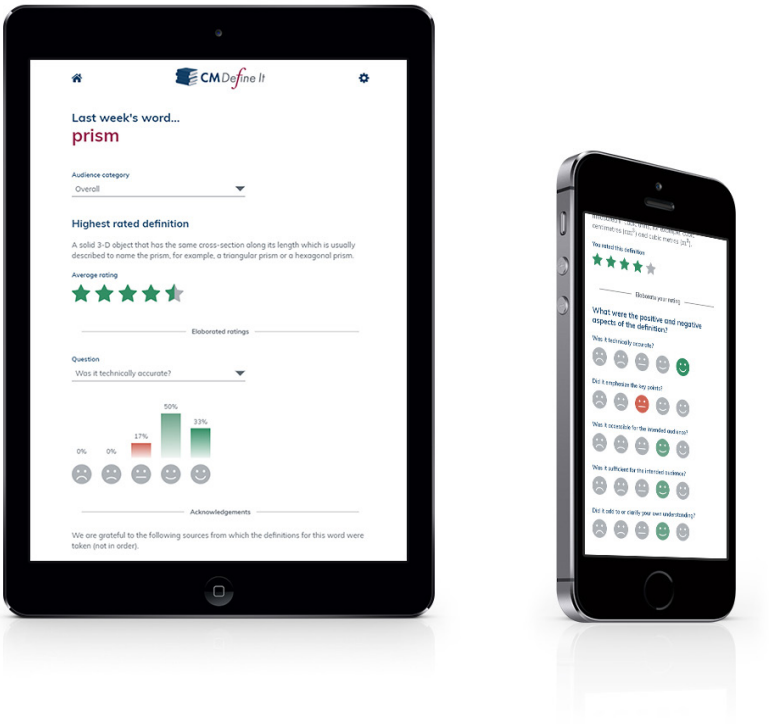
After selecting the category of learner, users rate the definitions. Each week, they are presented with a new word and up to five definitions of the word. These definitions have been taken, with permission, from a range of international sources that addressed novice, intermediate and advanced learners. These already-existing, reputable sources are acknowledged at the end of each definition screen.

Users are asked to give an overall rating of each definition on a scale of 1-5 stars with 1 star holding the lowest rating and 5 stars holding the highest. They are then asked to rate each definition on 5 additional criteria, including:

- how technically accurate the definition is;
- whether the definition emphasises key points;
- how accessible the definition is for the chosen audience;
- whether the definition is sufficient for the chosen audience; and
- whether the definition adds to or clarifies their own understanding of the word.

Free response boxes are also included if users want to provide additional feedback that is not captured in the above criteria.

After rating the definitions, users can view their previous responses and the last week's results. They can filter last week's results according to the audience category and the five criteria.



Pilot study

A pilot study is a small-scale run of the main research study, which investigates the procedure or research instrument intended to be used in the main study (Persaud, 2010). Pilot studies should be conducted with participants closely resembling the population targeted in the main study. They are very valuable when little is known about the topic of study or when using novel research instruments which may need validation and/or improvement. Pilot studies aim to discover issues that require adapting and solving prior to the main study, therefore they are an important step in the process of conducting research in novel fields, using un-tested instruments and procedures (Persaud, 2010).

The Cambridge Mathematics team conducted a pilot study of the *CM Define It* app. The aims of the pilot study were:

- to test the app for the first time,
- to assess its functionality,
- to assess user experience, and
- to use participants' feedback to improve the app prior to its launch.

E-mails and letters with details of the pilot study were sent out to academic staff and schools in the UK and internationally. The pilot began at the end of May 2019 and lasted for seven weeks. The link to access *CM Define It* was available on the *CM Define It* website and the app could be downloaded from Google Play and Microsoft stores.

Overall, 36 users signed up to the pilot study of the *CM Define It* app. The most common main professional role was "teacher" with 18 users identifying themselves as teachers. This was followed by "educational researcher" with 9, "other" with 5, "maths/science academic" with 2 users, "teacher educator" with 1 user and "resource developer" with 1 user. Of those who selected an additional professional role, "teacher educator" was the most common additional job role with 7 and "student" was the least common additional role with 3 users. English was the most common language with 22 users selecting English (UK) and 3 selecting English (US) as their first language. However, there were also a number of other languages selected, including: Spanish, Serbian, French, Polish, Afrikaans, English (other), Swahili, Turkish and Shona. Although the majority of users came from the UK (20), 12 other countries were also represented: China, Argentina, Serbia, Belgium, India, South Africa, Botswana, Kenya, Uruguay,

Zimababwe, United States and Malaysia. The majority of users registered on a browser (23), followed by Android (9), Windows (3) and unspecified (1). Out of the overall 53 responses across the 7 weeks, 19 responses were made with the *novice* category in mind, 19 were made with the *intermediate* category in mind and 15 were made with the *advanced* category in mind.

Ratings data suggested an increase in participation in week 3, as 15 participants rated the word of the week compared to 6 in week 1 and week 2. Following on, from week 4 the number of responses did not go beyond 6. Despite what appears to be a small sample, pilot studies are almost always based on a small number of participants (van Teijlingen & Hundley, 2001).

Since the sample size was very small there is no statistical power in the results (Jameson et al., 2019; van Teijlingen & Hundley, 2001) and the results were used as indicative findings only. Future reports on any larger-scale studies may include statistical analyses if appropriate.

The next section will describe how participants' feedback on the *CM Define It* app was collected and how it contributed to re-development of the app.

User experience: post-pilot evaluation

Participants in this pilot study engaged in natural use of the app (Rohrer, 2014). Natural use aims to understand users' behaviour as close to reality as possible and allows the researchers to study behaviours and/or attitudes that participants would show when actually using the product, potentially increasing the validity of the study (Rohrer, 2014).

After the seven-week-long pilot study, participants were contacted for feedback regarding their experiences of using the *CM Define It* app. When evaluating the Statistical Interactive Glossary (SIG), Haas et al. (2003) noted the following:

- the content of the presentation should be accurate;
- the glossary's interface must be usable;
- the glossary help must be effective;
- the presentation should be appealing;
- the presentation should be informative; and
- the information should be given in a way that the users can understand (Haas et al., 2003, p. 198).

Väänänen-Vainio-Mattila, Roto and Hassenzahl (2008) state that to develop strong user experience (UX), researchers and/or designers should understand the functionality and personal needs which motivate the use of the product (p. 3962). A good user-centred design process includes iterative design and evaluation of prototypes, as without evaluation it is highly impossible to improve the design or product (ISO, 1999, as cited in Väänänen-Vainio-Mattila et al., 2008). When designing the evaluation survey, these issues were considered. A SurveyMonkey questionnaire with 30 questions was developed in order to evaluate the *CM Define It* app. Of the 30 questions, 27 evaluated the app, its usability, functionality and user experience. Examples of questions in the survey include:

- *Did you subscribe to e-mails reminding you to rate the word of the week?*
- *Was the interface of the app user-friendly?*
- *Could the interface of the app be improved?*
- *Did the three categories of learners (novice, intermediate, advanced) adequately capture all of the possible groups that may be learning mathematics?*
- *What other groups would you like to see included or how would you improve the three existing categories?*
- *How would you improve the definition rating system?*
- *What additional features or functions could be incorporated that would make you continue engaging with the app?*

Participant feedback and responses

All feedback was considered and appropriate changes were made to the app. Any suggestions not incorporated at this stage were not included for specific reasons and will be re-considered for the Framework glossary. Table 1 illustrates the feedback given by pilot participants in the post-pilot survey and how their feedback was acted on.

Table 1: Summary of feedback from the pilot phase of the CM Define It app and the changes made to the app in response to the feedback

Question	Feedback	Response to feedback
How long did you engage in the *CM Define It* app for?	2 weeks – 16.67% 3 weeks – 33.33% 4 weeks – 50%	No response required.
Did you subscribe to e-mails reminding you to rate the word of the week?	Yes – 100%	No response required.
Was the reminder helpful?	Yes – 83.33% Somewhat – 16.67%	No response required.
Would a reminder on your phone make you want to continue using the *CM Define It* app?	Yes – 33.33% No – 66.67%	Push notification – in the new version of the app, participants have the option to receive push notifications reminding them to complete their ratings for the word of the week.
Was the interface of the app user-friendly?	Yes – 100%	No response required.
Could the interface of the app be improved?	Yes – 33.33% No – 66.67%	No response required.
How could the interface be improved?	"It only uses words, not EAL friendly. Needs more visuals."	Images and diagrams to explain key words – the team considered including images and diagrams to illustrate concepts. However, if we were to include images, they would need to accompany every key word for consistency. Due to time constraints, the team decided not to include images in the re-developed app. However, after providing ratings, participants are presented with a question asking if a diagram accompanying the definition would be useful. The team will consider whether diagrams and images will be useful to include in the Framework glossary.
Did you experience any accessibility issues when using the app?	No – 100%	No response required.

Question	Feedback	Response to feedback
If you experienced any accessibility issues when using the app, please give details of the difficulties you encountered.		No response required.
Did the three categories of learners (novice, intermediate, advanced) adequately capture all of the possible groups that may be learning mathematics?	Yes – 83.33% Somewhat – 16.67%	No response required.
What other groups would you like to see included or how would you improve the three existing categories?		No response required.
Would you have liked to be able to rate definitions for more than one category of learners?	Yes – 66.67% No – 33.33%	The re-developed app allows users to go back and rate the definitions for another category of learners.
Was the five star rating system an effective way of rating definitions?	Yes – 66.67% Somewhat – 16.67% No – 16.67%	See below.
How would you improve the definition rating system?	"Enable changing the overall rate after grading all criteria (currently I am trying to "fit" all component rates into a previously given final). Maybe these components could be listed before the final rate?" "Three star is enough".	Although the order of ratings remained the same, the app now allows users to change the rating they have already given to a word. Previous and new ratings are collected.

Question	Feedback	Response to feedback
If it was possible to provide ratings for words from previous weeks, would this make the app better?	Yes – 50% No – 50%	It was decided that if participants could change their previous responses, too much data would be available for analysis and the main aim of the app would be lost. The new version of the app allows participants to change the rating they have awarded for the current word of the week but not any earlier ones.
Did the criteria adequately capture everything that mathematical definitions could be judged on?	Yes – 83.33% Somewhat – 0% No – 16.67%	See below.
If you feel there are other criteria that could explain the ratings you awarded to definitions or have suggestions for improving the existing criteria, please provide them here.	"Does it communicate conceptually? Is it using words that are appropriate for the category of student? Is the definition simple and clear enough? Would it best be illustrated?"	It was decided that the suggested criteria are very similar to the current criteria therefore no response was taken.
Do you feel the process allowed you to express your opinion about the definitions?	Yes – 83.33% Somewhat – 16.67%	No response needed.
How could the process be improved?		No response needed.
How did you engage with the *CM Define It* app?	Web link – 83.33% Web link followed by the app downloaded on an Apple device – 16.67%	The *CM Define It* app is now available for download on Android, Apple and Microsoft devices.
Did you experience any difficulties with accessing and/or changing your personal information?	Yes – 16.67% No – 83.33%	No response needed.

Question	Feedback	Response to feedback
What difficulties did you experience and how could the Cambridge Mathematics team improve the experience of editing personal information?		No action could have been taken due to lack of response.
How engaged were you with one word release per week?	Highly engaged – 16.67% Engaged – 83.33%	No response needed.
What frequency would you prefer?	Two words per week – 16.67% One word per week – 66.67% Other – 16.67%	As the majority of survey respondents suggested that they preferred one word per week, this was kept in the re-developed app.
What kind of analytics, if any, would you be interested in seeing after rating definitions?	The most popular definition from current week – 16.67% The most popular definition from previous week – 50% The most popular definition for the week for particular groups of users (primary school teacher, secondary school teachers, assessment developers etc.) – 33.33% The number of users from different groups participating each week – 16.67%	As survey respondents could select multiple responses in this question, many selected more than one possibility. The team considered each of the suggested analytics and decided to keep the analytics simple. App users can see previous results for the overall submissions or for submissions by learner categories (novice, intermediate and advanced). It was deemed too complicated at this time to introduce analytics by learner categories and job roles.
What additional features or functions could be incorporated that would make you continue engaging with the app?	An interesting mathematical idea or fact – 66.67% An interesting idea or fun fact that is not mathematics-related – 16.67% Other – "diagrams to illustrate definitions"	All suggestions were considered by the team in the re-development of the *CM Define It* app. However, we decided that we did not want to introduce mathematical ideas or facts as we did not want to bias users against other ideas. Additionally, all of the suggestions would require time and resources, which at this moment we cannot commit to. Diagrams to illustrate definitions were considered and will be considered for the Framework glossary. For more detail, please see above.

Launch

The app was announced on a range of platforms, including but not limited to the Cambridge Mathematics newsletter and Twitter prior to its launch. We created a range of information sources, such as short and long text descriptions of the *CM Define It* app, PowerPoint presentations and videos, which were then sent to different audiences to communicate this new development. Some individuals and schools were already aware of the pilot stage, but the majority were not. The app was launched on 18th October 2019. Collected data will inform the Cambridge Mathematics team of what the education community values in good mathematical definitions and what we should consider when designing the Framework glossary.

References

Ball, D. L., & Bass, H. (2002). Toward a Practice-Based Theory of Mathematical Knowledge for Teaching. In E. Simmt & B. Davis (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, AB: Canadian Mathematics Education Study Group.

Blackwell, C. K., Lauricella, A. R., Wartella, E., Robb, M., & Schomburg, R. (2013). Adoption and use of technology in early education: The interplay of extrinsic barriers and teacher attitudes. *Computers Education, 69*, 310-319. doi: 10.1016/j.compedu.2013.07.024

Bouck, E. C., Satsangi, R., & Flanagan, S. (2016). Focus on inclusive education: Evaluating apps for students with disabilities: Supporting academic access and success. *Childhood Education, 92*(4), 324–328. doi: 10.1080/00094056.2016.1208014

Corpus. (n.d.). In *Oxford Dictionaries*. Retrieved September 2, 2019, from https://languages.oup.com/our-story/corpus

Fahim, M., Fat'hi, J., & Nourzadeh, S. (2011). Wordlists in language teaching and learning research. *International Journal of Linguistics, 3*(1), 1–13. doi: 10.5296/ijl.v3i1.1205

Gómez, P., & Cañadas, M. C. (2013). Development of a taxonomy for key terms in Mathematics Education and its use in a digital repository. *Library Philosophy and Practice (e-journal)*. 903. 1-9.

Gómez, P., Cañadas, M. C., Soler, C., & Restrepo, A. M. (2009). Funes: un repositorio digital para publicar y compartir documentos en Educación Matemática [Funes: a digital repository for publishing and sharing documents in mathematics education]. In M. C. Cañadas, J. M. Contreras, & A. B. Heredia (Eds.), *Jornadas de investigación en el aula de matemáticas:dimensión histórica, social y cultural de las matemáticas* (pp. 271-278). Granada: SAEM Thales y Departamento Didáctica de la Matemática de la Universidad de Granada.

Haas, S. W., Pattuelli, M. C., & Brown, R. T. (2003). Understanding statistical concepts and terms in context: The GovStat ontology and the Statistical Interactive Glossary. *Proceedings of the American Society for Information Science and Technology, 40*(1), 193-199. doi: 10.1002/meet.1450400124

Jameson, E., Horsman, R., Macey, D., Gould, T., Rushton, N., Rycroft-Smith, L., Majewska, D., Stevens, B., & McClure, L. (2019). *Ontology: Structure and meaning in the Cambridge Mathematics Framework*. Cambridge, UK: Cambridge Mathematics.

Jameson, E., Rycroft-Smith, L., McClure, L., Majewska, D., Horseman, R., Macey, D., Gould, T., Rushton, N., Knight, R., Fortin, L., Stevens, B. (2019). *Ground tests and test flights: Refinements from formative evaluation of the Cambridge Mathematics Framework (poster)*. Cambridge, UK: Cambridge Mathematics.

Monroe, E. E. (1996). Language and mathematics: A natural connection for achieving literacy. *Reading Horizons: A Journal of Literacy and Language Arts, 36*(5), 367 – 379.

Monroe, E. E., & Orme, M. P. (2002). Developing Mathematical Vocabulary. *Preventing School Failure: Alternative Education for Children and Youth, 46*(3), 139–142. doi: 10.1080/10459880209603359

Persaud, N. (2010). Pilot study. In N. J. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 1033). Thousand Oaks: SAGE. doi: 10.4135/9781412961288

Riccomini, P. J., Smith, G. W., Hughes, E. M., & Fries, K. M. (2015). The language of mathematics: the importance of teaching and learning mathematical vocabulary. *Reading & Writing Quarterly, 31*(3), 235–252. doi: 10.1080/10573569.2015.1030995

Rohrer, C. (2014). *When to use which user-experience research methods*. Nielsen Norman Group.

Väänänen-Vainio-Mattila, K., Roto, V., & Hassenzahl, M. (2008). Now let's do it in practice: User experience evaluation methods in product development. *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, 3961-3964. doi: 10.1145/1358628.1358967

van Teijlingen, E., & Hundley, V. (2001). The Importance of pilot studies. *Social Research Update*, 35. Retrieved from http://sru.soc.surrey.ac.uk/SRU35.html