Methodology

Building the research base

Others in this series

- Glossary App
- Formative evaluation
- Research-informed design

Author

Ellen Jameson

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.

Methodology: Building the research base

Related design principles

- Research-informed
- Transparency

Sections in this document

Introduction

Research is input, output, and process in the design of the Cambridge Mathematics Framework. We are developing the Framework so that its structure and content are transparently informed by research, can communicate some of the curriculum implications of research, and can contribute to research in the future. We find and synthesise research done by others, generate content informed by interpretation of that research, and evaluate our sources, content and practices as we go. We trial the ways people might use the Framework in pilot projects and communicate about our ongoing activities with audiences and collaborators from teaching, design, and research communities. Our priorities for research and research methods are shaped by our aims for the role of research in the Framework and our constraints in time, resources, and the availability of relevant research in different areas.

Our goals for transparency

On one hand, our work on the Cambridge Mathematics Framework is influenced by theoretical perspectives, international evidence, and empirical research. On the other hand, we are also influenced by the cultural and institutional contexts of the research, curricula and teaching practices that we and our collaborators are most familiar with from our own backgrounds. This means that various factors shape how we find existing research and interpret its implications for the design of the Framework. While we can try to survey areas of the literature from as many different contexts as possible, the Framework structure will still reflect what we are able to find and include, and the choices we make on that basis. This makes it essential for us to be as explicit and transparent about our design choices as we can so that future users of the Framework can make their own informed decisions.

In order to meet our goals for transparency, our work is guided by four main questions:

1. What can we say, as the Framework designers, about what causes our own attention and decision-making to go in certain directions and not others?
2. What literature review process should we follow that is systematic enough for us to identify the themes we should focus on, while remaining limited enough in scope for us to complete it for the essential core of the Framework?
3. How much should we report and how should it be linked to content and structure?
4. Eventually we will also be able to ask: *in what ways do people using the Framework refer to these explanations when making their own design decisions?*

Our goals for the Framework's research base

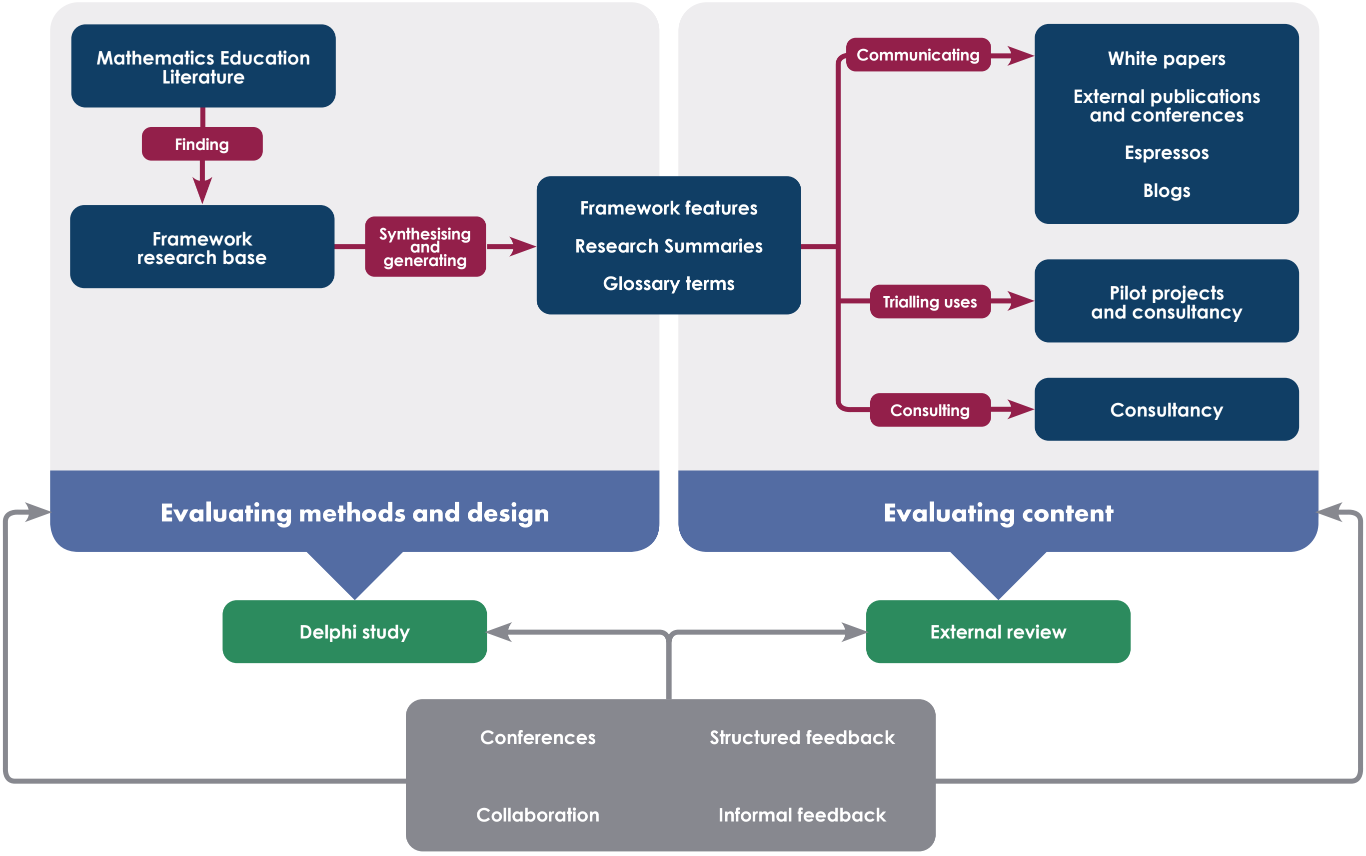
Our aim is for the research base to include key theoretical and empirical influences which are relevant to curriculum content and structure and which are generally considered to be of good quality within the mathematics education research community. The Framework should represent this research in a form that is a reasonable, useful, and approachable representation for our main user audiences – curriculum developers, resource developers, and teachers. We draw on a mixture of theoretical and empirical influences from the mathematics education research literature, as well as expert experience in research and, to a certain extent, mathematics teaching practice. Research is generally not prescriptive even when it is deliberately aligned with curriculum and policy questions. Therefore, our process of finding and interpreting the variety of sources we work with is an important part of how others will interpret the Framework, and becomes part of what we write about research in the Framework and related documentation.

Research in the Framework design cycle

As part of the overall process of developing the Cambridge Mathematics Framework, we are approaching research from several complementary angles. Over the course of our design cycle (see Figure 1), we are doing secondary research when we find and synthesise sources in the literature, and generate features from that synthesis. We do this to inform various specific topic areas within the Framework, to develop our methods for design and evaluation, to share what we find with a wider audience in Espressos or reports, or to build the theoretical basis for our design principles. When we conduct interviews and surveys, we are doing primary research. We do this to evaluate the structure and research basis for topic areas within the Framework, the design-in-progress of the Framework as a whole or the needs of potential Framework users. Communicating about evidence and design helps us to keep in touch with the academic research community about both approaches, as well as with communities of potential Framework users (see Figure 1). We are also structuring our current research activities with an eye towards laying the groundwork for future research, involving both Framework-aligned designs and the structure of the Framework itself.

CAMBRIDGE
√Mathematics

Figure 1: Research in the Framework design cycle: Finding, synthesising, generating, evaluating, trialling and communicating



The research methods we follow in each case result from a combination of considerations:

- The research available in the literature
- Our capacity to find, evaluate and synthesise relevant existing research
- Our criteria for judging existing research to be relevant
- Our perspectives on mathematics teaching and learning
- Our perspectives on evaluation and knowledge sharing within and between multiple communities in curriculum design and enactment
- Our perspectives on the role of research in design methodology
- Our perspectives on scenarios for Framework use and impact.

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

Boundaries of the research base

In broad strokes, the existing body of research in mathematics education can be characterised according to:

- the nature of the questions that tend to be asked (or not) about mathematics learning at different stages of schooling,
- the methods used (or not) to address them, and
- the degree to which experiments and outcomes may be shaped by existing curricula and pedagogical practices embedded in different sets of cultural and institutional contexts.

Participants on a Delphi panel of expert curriculum researchers and designers have contributed to the following sketch of the current state of the research literature on mathematics learning (2018) that is most relevant to the design of the Cambridge Mathematics Framework:

- Research with younger students is more likely to focus on their thinking and how they learn foundational mathematical ideas. Less of this type of research has been done on the mathematical ideas that older students are learning or building upon.
- Research with older students is more likely to focus on misconceptions or behaviour in specific interventions. There is a general gap in our picture of mathematics learning from year 8 until calculus, during which time a lot of conceptual development is happening very quickly.
- There is a general lack of longitudinal studies examining the longer-term implications of earlier mathematical experiences.
- There has been a great deal of theorising around the role of classroom communities of practice and learning environments in developing students' abilities to think mathematically, but while there is evidence that these are important considerations, much is still unknown about how this supports mathematical thinking.
- A lot of variation in individual learning trajectories which has been observed while looking for common patterns in learning has yet to be further explored.
- Domain-specific conceptual development is an important research gap.
- There is still a lot of ground to cover in characterising the links between teachers' content knowledge, pedagogical knowledge, beliefs and actions, and student learning.
- There have been long-standing calls for research on the affordances of new technologies with respect to the ordering of content, inclusion of content, and age of accessibility of content in curricula, but important gaps remain in this area.

Whether research has been published or not on a particular topic is the most fundamental factor determining how directly a given part of the Framework can be based on research. Beyond that, the nature and context of specific sources also plays a role. As a further example of the diversity of approaches and contexts represented in the literature, Kieran, Doorman & Ohtani (2015) describe reported overviews of the differences in philosophical perspectives and instructional goals and approaches, influenced by different lines of research in psychology.

> *"Ernest (1991) distinguishes four sets of issues related to one's philosophy of mathematics education: the philosophy of mathematics, the nature of learning, the aims of education, and the nature of teaching. In this regard, Burkhardt (2014) points out that different groups of people have different priorities with respect to curricular aims or goals in mathematics: "basic skills people", "mathematical literacy people", "technology people", and "investigation people". Likewise, Treffers (1987) distinguished four trends in instructional approaches to mathematics in terms of "horizontal" and "vertical" mathematization: mechanistic, empiricist, structuralist, and realistic, with each instructional approach drawing upon different psychological backgrounds - Gagne's cumulative learning for the mechanistic, Piaget's constructivism for the empiricist, Bruner's modes of representation for the structuralist, and Gestalt psychology for the realistic." (Kieran et al., 2015, p. 64).*

For us, this serves to illustrate not only the range of research foundations, designs, and objectives to be found in the literature, but also the range of reasonable interpretations that are possible based on relevance and implications for different audiences within education systems. From a framework design perspective, this wide range is one of the things that makes transparency a key design principle because it means that it is essential for us to explain why we have highlighted concepts and relationships between concepts in specific ways within the Framework.

Below, we describe our process for finding research sources and our two-stage process for evaluating sources for relevance and inclusion in the Framework.

Our capacity to find, evaluate and synthesise relevant existing research

A systematic literature review in every area of school mathematics learning from age 3-19 is not possible for us as a prerequisite for design; the relevant literature is vast. However, many projects have undertaken portions of such a review and we refer to their work whenever possible. For the sake of identifying important themes and findings across the breadth of the Framework we are following a semi-structured review process that includes keyword database search, purposive sampling according to syntheses, meta-analyses and recommendations from collaborators, and breadcrumb search starting from widely accepted texts such as recent research handbooks (Thomas & Harden, 2008). Similar to the process for thematic synthesis that Thomas and Harden describe, we make greater use of grey literature than a more structured review would, but we feel this is appropriate to our aim of broadening our initial coverage of the literature to determine what seems most essential to address in the Framework.

The Framework writing team has developed a set of practices and sensibilities for reviewing and synthesising research which they follow to write *Research Summaries* and to generate corresponding sets of features and glossary entries in the Framework (see Fig. 1). The format and scope of Research Summaries were developed as a way to standardise the process of synthesis for Framework writing, for the sake of the writers and eventually for the audience. Research Summaries are short papers which explain how research has informed the structure and content of a section of the Framework. They consist of:

- A topic which tells the story of a coherent collection of mathematical ideas in the Framework (with the size of the collection being appropriate to the topic but also usually within an agreed range of the number of ideas that will be highlighted),
- A literature review in which the writers aim to summarise as complete a range of important themes and findings from the literature as possible, even if complementary or apparently contradictory, including a justification for the framing of the collection itself,
- An interactive map of the collection of mathematical ideas from the Framework which are featured in the Research Summary, including the ability to view descriptions and contextual information for each idea, and the relationships between ideas which the author has focused on, and
- A section discussing implications of the research in the literature review for the content and structure that have been generated to form the collection of mathematical ideas featured in the Research Summary.

Each team member's way of reviewing and synthesising research for the Framework has converged over time in some important ways as the result of regular meetings and conversations to compare work on Research Summaries, resulting in documents which explain and show how research informs sections of the Framework in a reasonably consistent way.

When beginning a Research Summary, a member of the writing team might use questions similar to these to help provide initial structure to their literature search. They start with an initial idea of a topic or collection of related mathematical ideas:

1. What aspects of this topic have been prioritised and/or addressed in the literature?
2. What might this topic or collection of ideas lay the foundation for later on?
3. What goals for instruction have been identified in the literature?
4. Are there aspects which seem relevant to us for which little or no literature seems to exist in this field? In other fields?
5. Does empirical research on learning trajectories or progressions exist for this topic? Over what range?
6. Why should this topic or collection exist in this form? Is it contributing to making something explicit which is important but usually tacit for designers, teachers or researchers? Is it a topic for which the literature has identified important issues?
7. What scope of topic is meaningful to address in a single research summary? Should additional related Research Summaries be planned to expand on this topic appropriately?

Our criteria for judging existing research to be relevant

We consider the quality and context of research influences when evaluating relevance.

Source type and context

When considering source types and contexts, we use research influences from mainstream instructional contexts and also expert perspectives from research and practice that might extend beyond these contexts. If none of our influences come from outside existing curriculum design or practice, we would risk reproducing the current status quo. At the same time, there are at least two good reasons for considering research on mainstream approaches to be highly relevant. First, many of the short-term uses of the Framework in curriculum refinement, resource design or teaching would be taking place in typical education contexts for a given jurisdiction, shaped by existing learning and assessment objectives. Second, much of the available relevant research has taken place in mainstream instructional contexts and it is often not possible to compare outcomes of specific approaches outside of these contexts. We record source contexts according to the following categories: *journal article; conference paper; book chapter; book; textbook; white paper; expert collaboration; author direct experience; author synthesis*; or *other*. For each context we also record peer review status if this can be determined.

Most of our sources come from research handbooks that are widely cited and/or edited by established senior researchers in mathematics education, highly-ranked peer-reviewed academic journals, more specialised peer-reviewed academic journals with an excellent reputation in relevant subfields, or teacher-facing journals providing a high-level overview of research implications for curriculum and pedagogy. Our team examines whether the authors of most of our sources, both significant and supporting influences, have an established reputation for trustworthy and relevant work in their field. If not, there are still plenty of reasons this source may still be worth considering for inclusion with further expert review. See Fig. 2 below for a detailed flowchart describing our evaluation process.

We also categorise the type of each source we review according to the focus of the work presented by the author (e.g. theoretical, empirical, literature review, meta-analysis) so that we can evaluate it and use each source appropriately. This also allows us to characterise further the influences that make up our research base as a whole. These categories are shown in Figure 2 and explained below.

An explanation of source types

We initially formed these categories after discussing the nature of the sources we had been finding in our exploratory work and the types of sources we expected to find based on past experience. We have paid special attention to the relationships between empirical data, theory and philosophy, and the forms these can take in the mathematics education literature. We conduct occasional inter-rater reliability exercises to revisit the categories and determine whether they are being applied consistently and meaningfully or whether they might need to be modified.

Meta-analysis: A paper that describes itself as a meta-analysis will be categorised as one if it reports methods which are appropriate for meeting that claim:

- Systematic criteria for literature search and inclusion
- Systematic synthesis of findings that were reported in all the papers which met the search and inclusion criteria, usually based on statistical analysis of standardised effect sizes to produce a statistical result, such as an estimate of an effect (Kulik & Kulik, 1989).

A poor meta-analysis would not be included as even a secondary Framework influence except under specific circumstances requiring special explanation and reviewer approval.

Figure 2: A flowchart showing research base pre-review inclusion criteria by source type

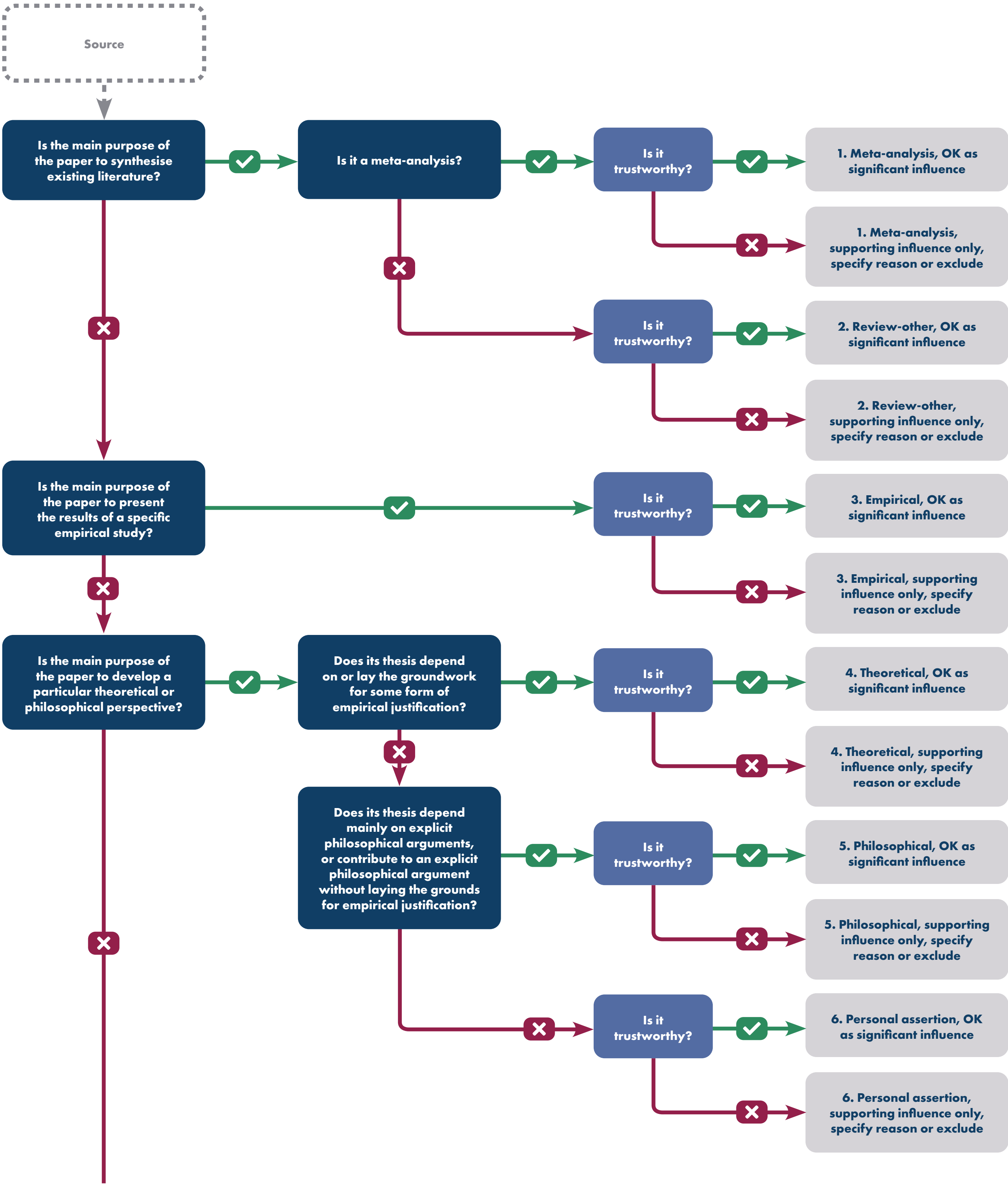
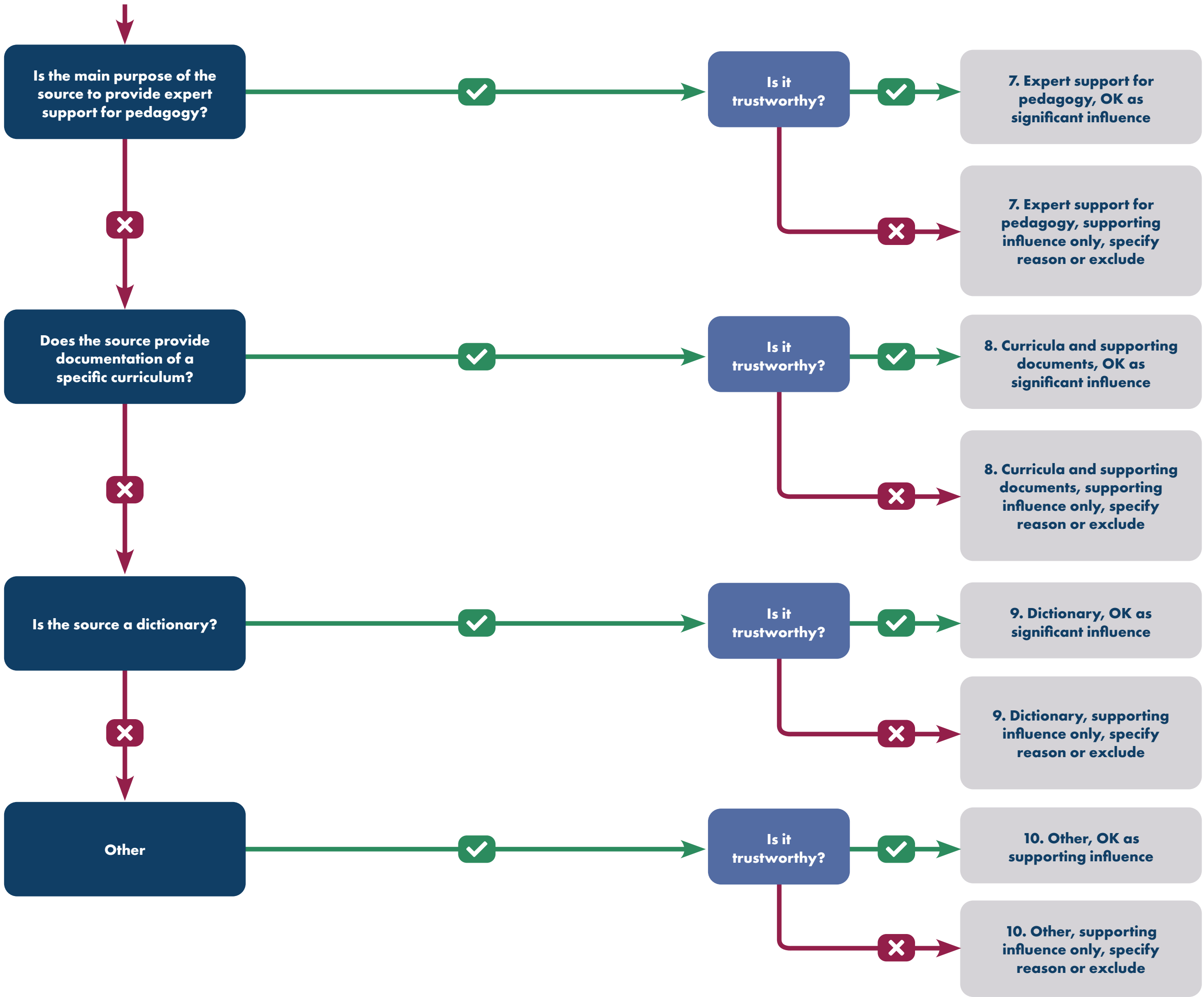


Figure continues on next page



Review-other: We put a paper in the *review – other* category if it is a literature review/research synthesis that presents multiple studies and their findings in a specific field (*e.g.* children's ability to learn long multiplication), but is not a meta-analysis. Since most empirical papers will include this type of review to some extent in their introduction-background-context sections, we only categorise a study as a *review-other* if it is focused on explaining different studies throughout the whole paper and not just in its introduction.

Empirical: Papers in this category *primarily* present the methodology and results of a particular empirical study (as opposed to primarily drawing conclusions about several studies reviewed together, or using an empirical review to refine theory).

Theoretical: We will put a paper in our theoretical category if it *primarily* focuses on laying out or developing theory in detail. It may even suggest what could be tested, but does not mainly focus on reporting empirical observations related to the theory. This category has been (rightly) difficult to distinguish from philosophy, but we still feel it is relevant to do so for some sources. Our reasoning is described in more detail below.

Philosophical: We label a paper as philosophical if it is *primarily* concerned with developing or discussing philosophical arguments rather than theoretical arguments or empirical data. We describe arguments as philosophical if they do not primarily depend on empirical data and are not necessarily testable, though empirical studies may still be part of the discussion. Our reasoning is described in more detail below.

Personal assertion: The source (or cited portion of the source) primarily lays out a personal perspective without reliance on support from defined philosophical arguments, theoretical constructs, or empirical evidence.

Expert support for pedagogy: Pedagogical support materials written by established researchers in mathematics education or by mathematicians. Such a source might elaborate on content knowledge or pedagogical content knowledge, or it might provide an annotated set of suggested activities.

Curricula and supporting documents: In this case, we take curriculum to mean “official documentation that sets out the elements of a course of study” (“Curriculum,” 2019). This can include national curricula and supporting documentation.

Dictionary: Dictionaries primarily provide definitions of words, often based on widely agreed convention and linguistic research.

Distinguishing between source types: philosophy, theory, and empirical research

In addition to helping us decide how to evaluate different sources appropriately, another purpose of creating separate empirical, theoretical, and philosophical source categories is so that we can better examine the types of influences underlying different areas of the Framework. However, by doing this we do not mean to imply that these things are ever wholly separate in the literature. The nature of theory and its relationship to philosophy and to empirical research in the social sciences, including education, is defined in wide-ranging ways when it is defined at all (Abend, 2008). Philosophy, theory and empirical inquiry are linked naturally in any research context, with philosophy providing the perspective that lays out the nature of what can be known about the object of the theory and how we might come to know it by empirical or other means. As a result, most if not all sources in the literature will discuss some blend of philosophy, theory and empirical inquiry either implicitly or explicitly.

In the education research literature in particular, the interdisciplinary and applied nature of the field and its societal roles contribute to greater difficulty in distinguishing between theory and philosophy (Siegel, Phillips, & Callan, 2018). The Cambridge Dictionary defines philosophy as "the use of reason in understanding such things as the nature of the real world and existence, the use and limits of knowledge, and the principles of moral judgment" and "a particular system of beliefs, values, and principles" ("Philosophy," 2019). We therefore distinguish arguments which are primarily philosophical from those which are primarily theoretical in part by their focus on the logical implications of prior philosophical arguments.

Evaluating source quality

In terms of characterising the research base, our first-pass evaluation of sources is more focused on context, but for the purpose of deciding whether to include a source, we examine its quality and relevance according to the type of source it is. These considerations are particularly important when deciding whether to include any sources that do not come from established research handbooks or high-quality journals. Figure 2 represents these judgments as a determination of *trustworthiness*, and a source which is not deemed trustworthy may still be judged to have enough *authenticity* to be included for some specific, well-delineated purpose. Both of these terms are described below along with their importance in our evaluation of sources.

These terms, trustworthiness and authenticity, are important ideas in the framework for constructivist inquiry laid out by Guba and Lincoln (1989, 1994), and are highly relevant to the sources we consider for the research base. They defined *trustworthiness* as a concept for meaningfully evaluating the quality of research within the constructivist paradigm that parallels the concept of *rigour*, which describes the quality of research in the positivist paradigm. Guba and Lincoln characterise trustworthiness in terms of a set of criteria which parallel conventional criteria for rigour, and which have been widely adopted for evaluating the quality of qualitative research (described in Table 1).

Table 1: Parallel criteria for evaluating research methods, based on Guba and Lincoln (1989).

Rigour (positivist)	Trustworthiness (constructivist)
Internal validity: a study has reasonably identified what it is measuring	Credibility: "the constructed realities of respondents (or stakeholders)" are likely to be well-represented by the researchers based on steps they have taken to verify this (p. 237).
External validity: the results of the study can be applied to other contexts	Transferability: Enough contextualised information has been provided about findings that others may recognise a way to apply or adapt these findings in their own context.
Reliability: replicability, as a measure of "consistency, predictability, dependability, stability, and/or accuracy" (p. 235)	Dependability: The research methods have been documented and conducted in a way that another study in a sufficiently similar context ought to yield similar findings.
Objectivity: appropriate minimisation and treatment of bias are expected to yield relatively unbiased data.	Confirmability: Parties external to the study, like auditors or reviewers, have been or would be able to confirm that the links between the details in the data and the researchers' interpretations of the data are reasonable.

When evaluating source quality for the Framework, we engage with trustworthiness rather than rigour. This is partly because of the makeup of the relevant mathematics education literature we draw on (many sources which mainly report empirical research or theory development are taking qualitative or mixed-methods approaches), and partly because for purely quantitative research we would need our external reviewers to make a final determination of rigour in questionable cases. In the Framework writing context, we can only apply these trustworthiness criteria as far as our backgrounds and experience with the literature will allow. We collaborate with external reviewers who have backgrounds in appropriate areas of mathematics education research, who use their own judgement in making their own evaluation of significant sources.

Trustworthiness criteria focus on a source's methods, but we have found that sometimes we have additional reasons for including a source as an influence in the research base. As shown in Figure 2 our inclusion criteria allow us to make the decision to continue to include a source even if we do not judge it to meet the criteria for methodological trustworthiness. If we included such a source, we would be doing so based on judgment of the source's *authenticity*, based on knowledge of the context and conduct of the research and the impact it may already have had in mathematics education. If a source was included as an influence solely for this reason, its influence would be carefully explained and qualified by the writing team and specifically evaluated in external review.

The concept of *authenticity* was developed as a means of expressing the quality of research in terms of its capacity to make meaningful, practical contributions to a community – starting with the experiences of the research participants themselves. Guba and Lincoln (1989, 1994) developed criteria for authenticity based on the notion that "[o]utcome, product, and negotiation criteria" are as important as method "in judging a given inquiry" from a constructivist perspective (1989, p. 245-250), expressed here in terms of our mathematics education context:

- *Fairness* in the source's representation of the constructions and value systems central to the inquiry, examining the role of stakeholders in building that representation and in determining implications
- *Ontological authenticity:* Stakeholders have been able to learn something of real value to them about the nature of their own mathematics education context from the source, whether by direct participation in a study or vicariously through reading about it.
- *Educative authenticity*: Stakeholders have been able to learn something of real value to them about the perspectives, knowledge and/or values of stakeholders in other communities.
- *Catalytic authenticity*: The source has been considered important enough by some part of the mathematics education community to put some of its implications or recommendations into action.
- *Tactical authenticity*: The source provides meaningful and useful tools – whether information, perspectives, arguments, design principles, interventions, etc. – that enable stakeholders to act in their communities.

To the best of our ability we consider the possible reasons for authenticity described above, and refer to the following general criteria for trustworthiness when evaluating sources. External reviewers then add their experience to the process of making a final evaluation for inclusion.

General criteria for trustworthiness:

- Is the research question(s) appropriate to the topic(s) and perspective(s) described in the paper, and does the paper adequately support development of the question(s)?
- Does the paper select and integrate appropriate theoretical models?
- Are the methods appropriate to the research question?
- Do the results and conclusions seem appropriate based on the topic(s), perspective(s), question(s), and methods above?
- Does any other work cast significant doubt on the results, conclusions, or justifications reported in the source?

Expert reviewers evaluate the first-pass outputs from our literature review in the form of Research Summaries. In these, we report our sources as either *significant* influences, which have supported specific decisions about structure or content, or *supporting* influences, which have been used to examine or support further the significant influences. We ask participants in the expert review process to evaluate significant influences according to the same general criteria as they would for peer review in general. They may tell us to:

- Include
- Include with specific caveats
- Exclude

Our perspectives on mathematics teaching and learning

We are designing the Framework to focus on flexibly identifying, describing, and connecting mathematical ideas which are important for learning mathematics. We emphasise flexibility because there is no absolute optimal order of mathematical experiences that will be best for every learner, and because there is also no single order that will be possible to offer in every jurisdiction. The dynamic visualisation of ideas and connections in the Framework is meant to allow designers and teachers to use the Framework to explore consequences of choices that they are making, for reasons that they may or may not have the power to change in order to do what seems best as they go.

Some of these choices may involve designers' and teachers' own perspectives on mathematics teaching and learning, as well as the perspective adopted by the curricula they're working within. Many of the learning theories which influence our own perspective are constructivist, and certainly this has had an impact on the fundamental structure of the Framework. Our decision to represent researcher and designer knowledge of student learning in the form of mathematical ideas building on one another in terms of student experiences and actions reflects our belief that this is an authentic means of working with this knowledge in ways that are likely to benefit students and teachers.

Constructivist theories of learning suggest that the ways individual students structure their understanding throughout the learning process can and will vary from one another and will not necessarily be directly visible to a teacher or researcher. This perspective carries implications for our use of research in the Framework; it affects the way we receive and interpret research about student learning. From our perspective a researcher in an empirical source must interpret what students understand, how that understanding changes over time, and why, and must adequately explain and justify their interpretations (see Evaluating source quality above).

Furthermore, constructivist theories of learning are not theories of teaching, and because we need the Framework to be compatible with a variety of different teaching perspectives and contexts, we have tried to represent mathematical ideas in a way that would be recognisable and useful from different teaching perspectives.

Our perspectives on evaluation and knowledge sharing

As we create the Framework and the tools and interfaces through which it can be used, we apply the ideas of trustworthiness and authenticity described above to our process of gathering feedback to evaluate and improve our design. As shown in Figure 1, we conduct primary research to gather and incorporate feedback at several points in our design cycle, and as we repeat stages of the cycle we make adjustments on the basis of that feedback.

We have developed our Framework evaluation methods with the goal that the Framework will be meaningful, useful, and used, in some form, by curriculum and resource designers, teacher educators, and teachers. Members of our core design team have experience in mathematics teaching, curriculum design, assessment design, textbook writing, resource design, education research, classroom action research, designing and delivering professional development, and information systems development for teachers. We work with a wider circle of collaborators which currently consists of mathematics education researchers and designers working in many contexts around the world, and pilot project collaborators who are willing to trial uses of the Framework in cases which represent the types of uses of the Framework we think could be widespread. As we proceed, this circle will widen further to include more collaborators from among the members of our user base.

Likewise, our evaluation methods are focused on face validation with our different prospective audiences. Using the Delphi expert survey method and interviews with researchers, expert review, case studies of pilot projects with designers, and surveys and interviews with teachers, we look at whether researchers, designers, and teachers can each find the meaning we intend to represent in the Framework, to what extent they agree with it according to their professional experience, and how they could find it useful.

Our perspectives on the role of research in design methodology

We draw on models for design processes in education that have been developed and refined within design research methodology in education for over twenty years (Cobb, Confrey, Lehrer, Schauble, & others, 2003; McKenney & Reeves, 2012). Particular aspects of design research that make some of its methods appropriate for us (apart from a general focus on design) include: linking specific design priorities and choices to theory; using initial design work to develop design principles that inform ongoing work; going through iterative cycles of design in which feedback on work in progress is incorporated into new design versions and practices; and participation in design by experts in multiple relevant communities (Barab & Squire, 2004; McKenney & Reeves, 2012; van den Akker, Gravemeijer, McKenney, & Nieveen, 2006).

We consider our process to be research-informed design because our aims differ in some respects from those of design researchers. For example, design methods are typically employed in primary research in education with the goal of producing or refining theories of learning. This is done on the basis of new data generated from direct implementation of research-based designs in classrooms or education systems (Barab & Squire, 2004). These designs and implementations are engineered so that together they may coherently contribute to theories of learning more broadly. In contrast, we rely on reviews of existing research to inform initial design choices, which we adapt and refine using feedback from expert evaluation. We aim to conduct and document our work in such a way that our design might later be able to contribute to primary research. However, for many areas of the Framework the closest we will come to generating data from the design before it is released will be through expert interviews and group surveys for face validation of the content and the structure of the Framework. In that sense, our methods in the current phase of Framework development are those that would be used in the beginning stages of design research.

Similar approaches have been described by framework development projects in other contexts. In a retrospective review of the standards writing process for the NCTM *Principles and Standards of School Mathematics* framework, the writers noted that a set of theoretical perspectives emerged as important influences over time and described their design work as "researchlike" because it also involved collecting, analysing, and incorporating feedback on work in progress (Ferrini-Mundy & Martin, 2003). Of work currently in progress, the UNESCO Institute for Statistics is currently developing the Reference List & Coding Scheme (RL&CS) framework, intended to provide rich qualitative support for mapping theory and curricula to assessment frameworks. This project's interpretive approach similarly required the designers to evaluate trustworthiness on the basis of practical value of the construct, as expressed through feedback on work in progress by expert members of the communities that would be making use of it (Cunningham, 2017).

Our perspectives on scenarios for Framework use and impact

The Framework is meant to be able to help designers realise the potential for some of the implications of existing research in mathematics education to be put into practice in the curriculum in a way that supports more coherent experiences of mathematics for students and teachers. The secondary research we do in our literature reviews, and our transparency in communicating research influences, can help to make the case for referring to the Framework in design projects and using it to examine curriculum, resource and assessment alignment. The primary research we do in our ongoing evaluation of the Framework can help it to take the form that will be most meaningful, useful and used by members of mathematics education communities. With both cases in mind, we hope that making mathematical ideas explicit which are the backbone of design in mathematics education will help them to be challenged, refined, and put to good use.

Bibliography

Abend, G. (2008). The meaning of 'theory.' *Sociological Theory, 26*(2), 173–199.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14.

Cobb, P., Confrey, J., Lehrer, R., Schauble, L., & others. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Cunningham, M. (2017). *Method for Developing an International Curriculum and Assessment Framework for Mathematics* (Internal Report) (pp. 1–15). UNESCO, Institute for Statistics (UIS). Retrieved from http://inprogressreflections.ibe-unesco.org/wp-content/uploads/2018/01/Math-Content-Reference-List-and-Coding-Scheme_Methods-Paper.pdf

Ferrini-Mundy, J., & Martin, W. G. (2003). Using Research in Policy Development: The Case of the National Council of Teachers of Mathematics' Principles and Standards for School Mathematics. In *A Research Companion to Principles and Standards for School Mathematics* (pp. 395–413). National Council of Teachers of Mathematics.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. SAGE Publications, Inc.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of Qualitative Research, 2*(163–194), 105.

Kieran, C., Doorman, M., & Ohtani, M. (2015). Frameworks and Principles for Task Design. In A. Watson & M. Ohtani (Eds.), *Task Design in Mathematics Education: an ICMI study 22* (pp. 19–41). Springer.

Kulik, J. A., & Kulik, C.-L. C. (1989). Meta-analysis in education. *International Journal of Educational Research, 13*(3), 221–340.

McIntosh, C. (2019). Philosophy. *Cambridge Advanced Learner's Dictionary* (4th ed.). Cambridge University Press. Retrieved from https://dictionary.cambridge.org/dictionary/english/philosophy

McKenney, S., & Reeves, T. C. (2012). *Conducting Educational Design Research*. Routledge.

Siegel, H., Phillips, D. C., & Callan, E. (2018). Philosophy of Education. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2018/entries/education-philosophy/

Stevenson, A. (2015). Curriculum. *Oxford Dictionary of English*. Retrieved from https://en.oxforddictionaries.com/definition/curriculum

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), 45.

van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational Design Research*. Routledge.