# TOWARDS A RESEARCH-BASED FRAMEWORK FOR STATISTICS EDUCATION

Darren Macey, Ellen Jameson, Rachael Horsman, Vinay Kathotia, Lynne McClure and Lucy Rycroft-Smith
University of Cambridge, UK
darren.macey@cambridgemaths.org

*Increasingly, calls are being made to raise the profile of statistical literacy in school mathematics curricula, with a growing body of research on the subject becoming available to curriculum designers and teachers. We are developing an evidence-informed framework for mathematics education for ages 3-19 to support curriculum design, teaching, assessment and professional development, as part of the Cambridge Mathematics project at the University of Cambridge. In this paper we discuss our design approach in statistics, describing how we developed a data structure that codifies a representation of statistical skills, concepts and processes, the key links between them, and links to other topics. We will describe the goals, compromises, assumptions and evidence base which have shaped our design, and we welcome feedback from the ICOTS community.*

## A FRAMEWORK FOR STATISTICS EDUCATION AT SCHOOL LEVEL

In order to support curriculum design and professional development Cambridge Mathematics aims to develop a broad map of the mathematical experiences that a school age student could reasonably expect to encounter. Within this landscape, the study of statistics and data handling presents a unique challenge to us as framework designers in representing students' development of technical skills and the judgment for using them, while capturing elements that distinguish this area of mathematics from others (Cobb & Moore, 1997).

Statistical literacy refers to "people's ability to interpret and critically evaluate statistical information, data-related arguments, or stochastic phenomena, and, when relevant, their ability to discuss or communicate their reactions to such statistical information…" (Gal, 2004); in other words, the ability to use statistical knowledge and techniques meaningfully and appropriately in context. There are consistent calls for statistics education to shift from a narrow focus on technique towards an emphasis on statistical literacy, reasoning and thinking more broadly, and to support this by encouraging students to engage with real data in familiar contexts in order to answer statistical questions (Franklin, 2007). Many researchers advocate engaging students in a process of data modelling (English, 2013; Lehrer & Romberg, 1996) beginning with their formative experiences in handling data and maintaining this approach as their skillset becomes more sophisticated.

To reflect this emerging goal, the Cambridge Mathematics Framework aims to support an increased focus on statistical literacy by explicitly including points of development of statistical literacy along with points of development in statistical techniques and clarifying key relationships between them. We expect this approach could be beneficial in jurisdictions where statistics curricula are defined as a list of mathematical techniques, leaving statistical literacy underrepresented (Department for Education, 2013). At the same time, we seek to highlight the representation of individual skills and concepts as they develop, supporting the design of curricula and programs of study without obscuring the data modelling process.

## DEVELOPING AN ONTOLOGY

The process of representing school mathematics involves consideration from multiple perspectives, including empirical research, curriculum development, teaching and learning, and professional mathematics. In the midst of the complexity of each perspective it is necessary to choose what to represent and how, a process requiring attention to both the structure of information that is included, and also what is left unrepresented. Our working definitions of the features of our Framework and the kinds of relationships we depict between them serve informally as an ontology for our representation (Fürst, Leclère, & Trichet, 2003), helping us to make more consistent decisions about how content needs to be represented in the Framework so that it meets our aims for showing meaningful connections.

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org 

Our early conception was of a mathematical landscape populated with ideas and experiences, with the Framework as something akin to a map showing the possible routes between these "waypoints". To this end the Framework is expressed in a graph database in which each waypoint is represented by a node and each connection by an edge. National Research Council (1990) guided the content focus of our early work alongside developmental ideas proposed by Bruner (1966) and Freudenthal (1983), establishing an initial design principle of generating connections from early, informal student experiences toward more advanced and formal conceptual work later in education. Originally we defined a waypoint as "a discrete location on a defined trajectory at which a learner's understanding of mathematical structure is added to or adapted", while the connections were simply directed or undirected.

Through an iterative design process similar to the process of ontology refinement described by Fürst et al (2003) we generated content and continuously reviewed the ontological structure among ourselves and with researchers in specific subdomains. Informed by the work of Tall (2013), Michener (1978a) and Hiebert & Carpenter (1992) we refined our ontology to include more types of waypoint and defined types for directed connections (Fig. 1).

| **Waypoint type** | **Purpose** |
|---|---|
| Standard | Learners acquire knowledge, familiarity or expertise |
| Exploratory | Set up useful intuitions; motivate basic definitions and results; establish elementary, but important, properties of concepts and examples |
| Meta | Define an approach to learning; stretches across subdomains of content |
| Gateway | Bring together pathways; check understanding |
| **Connection type** | |
| Conceptual development | A pathway through waypoints along which a concept develops |
| Use of process, skill, procedure | A connection where the content of one waypoints is used or applied at another |
| Undirected | A non-hierarchical relationship between content |

Figure 1. Provisional waypoint and connection types

## DESIGN IMPLICATIONS FOR STATISTICS

Meta-waypoints were specifically created to represent the cyclic approach to the application of statistics in data modelling, in the context of the individual skills available to students at various stages of development. These nodes collect together groups of waypoints and describe how the technical content interacts with overarching statistical principles.

In the example below, the meta-waypoint 'early data modelling' describes the initial experience students have with the modelling cycle. It encapsulates the modelling cycle within the context of the limited set of skills available to students at an early stage of learning mathematics while also defining pathways for the development of concepts. At this early stage, students begin representing data by grouping physical objects. The edges show the development of related concepts and their use as students begin to interpret their data, eventually creating iconic representations. At later stages of data modelling, (not shown in the example), students develop iconic representations into a broader representational toolset that encompasses one-to-many representations and representations that use length to represent frequency.

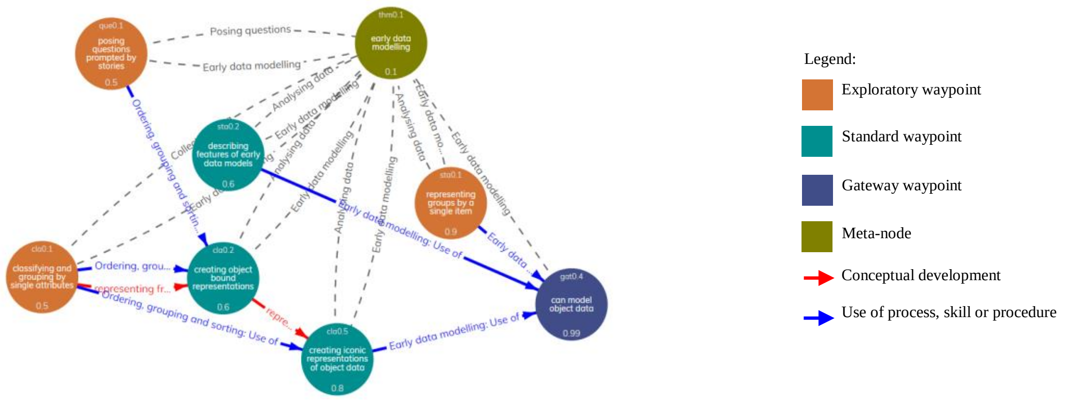


Figure 2. Early data modelling waypoints

The edges connecting these waypoints to the meta-waypoint describe the aspect of the statistical problem solving process attended to, splitting the statistical content into four stages.

| **GAISE report stages (Franklin, 2007)** | **Equivalent Cambridge Mathematics Framework stages** |
|---|---|
| Formulate questions | Posing questions |
| Collect data | Collecting Data |
| Analyze Data | Analysing data |
| Interpret results. | Interpreting results |

Figure 3. Stages of the statistical problem solving process

Meta-waypoints help to limit the complexity of the connections captured by the Framework: without them, a waypoint related to analysis of data would potentially require a direct link to every node related to the construction of graphical representations of data. These many-to-many connections are now mediated by a single meta-waypoint.

The remaining edges shown in the example connect standard and exploratory waypoints to each other, demonstrating hypothesised pathways for either conceptual development or use of a previously developed mathematical process, skill or procedure. These hypothesised pathways are established through a synthesis of research, expert opinion and practical classroom experience.

The subsequent meta-waypoint 'simple data modelling' (not shown in Fig. 2) collects a larger set of waypoints covering statistical content that students may reasonably be expected to encounter during the early primary years as they begin to explore methods of collecting, representing and interpreting data. As a result, two very different views are accessible to curriculum developers and classroom practitioners using the Framework. From one perspective using a meta-waypoint (Fig. 2), users can view the collection of mathematical techniques and experiences to be explored at various stages of education, and filter these based on which stage of the statistical problem solving cycle they attend to. From another perspective, users can follow a progression for the development of a particular mathematical theme (such as 'representing frequency by length') from its introduction to its culmination (Fig. 4).

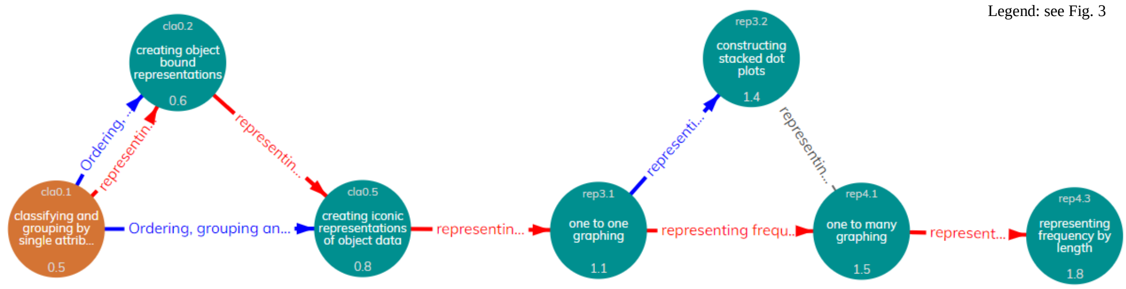


Figure 4. Hypothesised pathway of concept development for 'representing frequency by length'

Similarly, for teachers who want to attend to the statistical cycle, one view of the content allows them to identify the various techniques and skills that need to be rehearsed over time; and, for those aiming to move away from a focus on procedural development, an alternative view can show how the statistical cycle integrates with a more content-driven curriculum.

## UNDER REVIEW

The research base of the Framework is represented by an additional set of nodes in the graph database. It contains sources such as research papers and books alongside expert opinion. Because the scale of the project is large, we have taken a pragmatic approach to research coverage: placing greater initial emphasis on content specific research handbooks, expert recommendations and high profile reports. Known biases and other decision factors are discussed within the Framework in research summaries (documents detailing our synthesis of the research used) and research nodes.

An essential part of the process of validating the Framework is to open the content up for review, helping to address gaps in the research base and biases. The review process is ongoing throughout the life of the project, with the community of reviewers expected to expand beyond a limited initial pool of experts to encompass a wider community that incorporates the full range of potential future users of the Framework.

## BROADER GOALS

In this paper we have highlighted the process of designing a portion of a framework for mapping the skills, processes and techniques that make up the statistical knowledge base we believe students should experience in their journey through education from age 3 to 19. We hope that the Framework will assist teachers, curriculum designers and policy makers in developing a useable understanding of research.

## REFERENCES

Bruner, J. S. (1966). *Toward a Theory of Instruction*. Harvard University Press.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *The American Mathematical Monthly*, *104*(9), 801–823.

Department for Education. (2013, September). The national curriculum in England: Key stages 3 and 4 framework document.

English, L. D. (2013). Reconceptualizing Statistical Learning in the Early Years. In *Reconceptualizing Early Mathematics Learning* (pp. 67–82). Springer, Dordrecht.

Franklin, C. A. (Ed.). (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: a pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association.

Freudenthal, H. (1983). *Didactical Phenomenology of Mathematical Structures* (Vol. 1). Springer Netherlands.

Fürst, F., Leclère, M., & Trichet, F. (2003). Ontological engineering and mathematical knowledge management: A formalization of projective geometry. *Annals of Mathematics and Artificial Intelligence*, *38*(1), 65–89.

Gal, I. (2004). Statistical Literacy: Meanings, Components, Responsibilities. In *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Springer.

Hiebert, J., & Carpenter, T. P. (1992). Learning and Teaching with Understanding. In D. Grouws (Ed.), *Handbook of Research on Mathematics Learning and Teaching* (pp. 65–97). New York: Macmillan Publishing Company.

Lehrer, R., & Romberg, T. (1996). Exploring Children’s Data Modeling. *Cognition and Instruction*, *14*(1), 69–108.

Michener, E. R. (1978). *The structure of mathematical knowledge* (Technical Report No. 472). MIT Artificial Intelligence Laboratory.

National Research Council. (1990). *On the Shoulders of Giants: New Approaches to Numeracy*. Washington , DC: The National Academies Press.

Tall, D. (2013). *How Humans Learn to Think Mathematically: Exploring the Three Worlds of Mathematics*. Cambridge: Cambridge University Press.