Investigating changes in high-stakes mathematics examinations: a discursive approach

Candia Morgan & Anna Sfard

Published online: 06 Jul 2016.

Submit your article to this journal ⌴

Article views: 15

View related articles ⌴

View Crossmark data ⌴

Routledge
Taylor & Francis Group

🔓 OPEN ACCESS

Investigating changes in high-stakes mathematics examinations: a discursive approach

Candia Morgan[a] and Anna Sfard[b]

[a]Department of Curriculum Pedagogy and Assessment, University College London Institute of Education, UK;
[b]Department of Education, University of Haifa, Israel

ABSTRACT

This article focuses on the theoretical-methodological question of how to identify reform-induced changes in school mathematics. The issue arose in our project The Evolution of the Discourse of School Mathematics (EDSM), in which we studied transformations in high-stakes examinations taken by students in England at the end of compulsory schooling. We have adopted a conceptualisation that draws on social semiotics and on a communicational approach, according to which school mathematics can be thought of as a discourse. Methods of comparing examinations of different years developed on the basis of this definition enable identification of subtle disparities that are nevertheless significant enough to make an important difference in students' vision of mathematics, in their performance and, eventually, in their ability to cope with problems that can benefit from the use of mathematics. In this article, we present these methods and argue that they have wider application for comparative studies of school mathematics.

In the last decades, reforms in mathematics teaching and learning have been sweeping the world. The resulting changes express themselves in all aspects of school mathematics, with assessment procedures being in many cases among the first to be affected. For reformers, high-stakes assessment is one of the most effective vehicles for bringing curriculum change to schools. In this article we attend to the methodological question of how to investigate *the evolution of examinations* administered to students and hence to gain insight into the effects of policy and curriculum changes. The method proposed in this article has been developed in the attempt to study changes that have taken place during the last three decades in the high-stakes examinations taken by students in England at the end of compulsory schooling.

Obviously, we are not the first to undertake this kind of endeavour. The need for sensitive research methods that would allow identification of potentially consequential differences between different versions of school mathematics has been widely recognised, especially in the context of comparisons between countries, between 'reform' and 'traditional' versions within a single country, or between versions of school mathematics

offered to students from different social groups. The subject has been studied through the prisms of policy and curriculum documents (Hodgen, Marks, & Pepper, 2013; Smith & Morgan, 2016), textbooks (Dowling, 1998; Haggarty & Pepin, 2002; Herbel-Eisenmann & Wagner, 2007), classroom interactions (Andrews & Sayers, 2005; Clarke, Keitel, & Shimizu, 2006; O'Halloran, 2004) and even examinations (Britton & Raizen, 1996). Our project, however, has two features that set it apart from the majority of previous attempts. First, whereas numerous studies look for dissimilarities in various aspects of *different, usually more or less concurrent, instances* of school mathematics, there is a paucity of research on *change that happens over time*. A rare exception, Kilpatrick's (2014) recent historical review of changes in textbooks, focuses primarily on their form and function for teachers rather than on their relationship to student experience. Second, the method presented here is firmly grounded in the *vision of mathematics as a discourse*, an assumption that, although tacitly present in some previous comparative studies (Dowling, 1998; Herbel-Eisenmann & Wagner, 2007; Peled-Elhanan, 2015), has been an explicit basis for subsequent research in only a few of them (Newton, 2012; Park, 2016). As will be argued in this article, considering mathematics to be a discourse has significant implications not only for how change in examinations is investigated and for how informative and useful the results are, but also for what we consider as important and valuable in the process of learning and in its outcomes.

Our method of studying change in mathematics examinations over time and the underlying principles of this method are described in this article in general terms, although a number of brief illustrative examples have been added to make some of its aspects more concrete. More detailed presentations of various parts of our analytic scheme along with instantiations of its application and outcomes can be found in other articles in this special issue (Morgan; Morgan and Tang).

1. The context: the *why* and *how* of studying change in high-stakes examinations over time

In this section, we explain the origins of our method: first, outlining the context that gave rise to our interest in the evolution of high-stakes examination; second, listing the tasks we had to perform in order to construct the method; and third, presenting the rationale for choosing a discursive conceptualisation and method.

1.1. Why study changes in high-stakes examinations?

In recent years, the claim that "standards" of school mathematics "have fallen" has been a common motif in popular media and in the discourse of politicians in many countries around the world. Large-scale assessments and international comparisons such as TIMSS and PISA fuel this suggestion with metrics of decline or improvement in students' relative performance. However, the evidence and arguments presented in the media and mainstream policy documents generally lack nuance or deep understanding of issues involved in making such comparisons. Taking examples from the public media in England, from government press releases and from a substantial report produced by an influential UK centre-right think tank (Reform, 2005) we find the following types of argument:

- Slippage from writing about fall in England's position in the international ranking of PISA results to making claims about fall in standards (McSmith, 2007)
- A claim by a UK government minister that a fall in the success rate in exams will "inspire confidence" and provide a more accurate picture of standards (Paton, 2014)
- A statement by the same government minister that (following a change in policy about early entry to examinations) a rise in exam success rate is evidence of a rise in standards (Department for Education, 2014)
- Poll results showing public perception of fall in standards presented as evidence of the existence of a fall (Reform, 2005)
- The fact that universities have adapted their curricula, including making remedial provision for first year undergraduates, taken as evidence of a fall in standards of school education without taking into account massive demographic changes in the population entering university (Reform, 2005)

It could hardly be otherwise, considering that the key words of these debates are rarely explained in operational terms. This is true even with regard to the very notion of *standard*s and to the term *falling* as featured in the slogan "standards are falling". Policy decisions, rather than being informed by detailed analysis of all the available data, are justified with the crude and poorly understood tool of national and international assessments. As a result, the use of comparative data in policymaking, in England as in many other countries, tends to reflect "long-existing trends and characteristics of each national public debate in education" rather than supporting genuine attempts to identify, and cope with, existing problems (Pons, 2012, p. 213).

Policymakers need to know much more than statistically processed results of testing. If these debates are to be useful, sensitive, clearly defined methods for identifying changes must be used. Our research project The Evolution of the Discourse of School Mathematics (EDSM), was initiated with these needs in mind. Our first task in this project was to devise methods of study that would provide a sound basis for the evaluation of reform-induced changes and for subsequent decision making.

1.2. Designing a method for studying changes in examinations: a 'To do' list

The scarcity of proper evidence for the various claims about the nature and significance of reform-induced change is understandable, considering the challenges faced by those attempting to identify such changes. While launching our project, we faced three tasks related to data collection and data analysis. In this section, we describe these tasks and discuss their complexities (their implementation will be presented in section 2 of this article).

Our first task (Task 1) was to *construct the appropriate database.* Before we could address it, we had to answer the question of *what kind of data* would support refined, disciplined analyses of changes over time. An obvious response would be to turn to curriculum documents and syllabuses, which together give an overview of mathematical topics expected to be learned and of general approaches to teaching these topics. However, these types of data allow only broad comparisons and provide little insight into the detail of mathematical activity. To get a more useful picture of relevant changes, it would probably be best to study classrooms. However, although major studies such as

the TIMSS video project offer sufficient data to make comparisons across classrooms in different countries, they do not provide a *temporal* spread of the data that would make it possible to investigate the *processes of change*. We turned to examination papers as the best available archival window into reform-induced transformations. Our choice of public examinations taken at the end of compulsory schooling (since 1988 the General Certificate of Secondary Education [GCSE], prior to that the General Certificate of Education Ordinary Level [GCE O Level] and Certificate of Secondary Education [CSE]) seemed particularly appropriate, as they constitute a good reflection of mathematics discourse as practiced in schools in England. Indeed, the symbiotic relationship between assessment, on the one hand, and curriculum and pedagogy on the other is well documented (e.g. Barnes, Clarke, & Stephens, 2000; Broadfoot, 1996). The impact of GCSE examinations may be especially strong, considering the high stakes they have not only for students but also, through their use as accountability measures, for teachers and schools.

After deciding on the proper kind of data, Task 1 shrank to answering the question of *how to extract an appropriate subset from the large bulk of archival materials*. This subset had to be small enough to be manageable, but rich enough to provide us with a reliable image of the process of change.

Task 2 was related to the question of *how to analyse the thus collected data*. We were particularly concerned with the *sensitivity* of the method and its *operationality*. Consider, for instance, the questions from 1957, 1980, 1991 and 2011 examinations displayed in Figures 1, 2, 3 and 4, respectively (note that we do not treat these questions as capable of providing evidence of any general change in curriculum or examinations; they have been chosen solely as a vehicle to illustrate the kinds of differences we are interested in and to motivate the development of our analytic method). All the items can be roughly categorised within the broad topic of mensuration and, as such, are sufficiently similar to justify comparison. At the same time, they are clearly different. Traditionally, they would be described as showing differences either in their *mathematical content* or in the *level of difficulty* or in both. But is this characterisation informative enough? Does the gap between those who are able to answer some of these questions but not others mean that examinees merely arrived at different levels of mastery in *the same types of mathematical activities*? Taking a closer look, we begin noticing that there are many other differences we could consider. For instance, the four questions differ in their use of diagrams (absent in Figures 1 and 2), in the complexity of language (consider the third sentence in Figure 1 or part (iii) in Figure 2 compared to any of the sentences in Figure 4), in the way non-mathematical context is implicated, etcetera, and they also vary in the guidance and support provided to students, and thus in the opportunities

---

A uniform circular metal disc of diameter 15in. weighs 12lb. For the purpose for which it was intended its diameter must be $14\frac{3}{4}$ in. and its weight 11lb. When its diameter is reduced to the required size it is still too heavy and it is proposed to reduce the disc to the required weight of 11lb. by drilling through it a number of $\frac{1}{4}$ in. diameter holes. How many such holes must be drilled?

Figure 1. University of London, GCE Ordinary Level, Pure Mathematics (a) Arithmetic and Trigonometry, Summer 1957, question 8.

(i) Write down an expression, in terms of π, for the volume of a solid cylinder of radius 0.4 cm and length one centimetre.

(ii) A copper tube is in the form of a hollow cylinder. Its external radius is 0·4 cm and its internal radius is 0.3 cm. Taking π to be $\frac{22}{7}$, show that the volume of copper in a piece of the tube one **metre** long, is 22 cm^3.

(iii) Express 22 cm^3 in cubic metres, giving your answer in the form $A \times 10''$, where A is a number between 1 and 10 and n is a negative whole number.

(iv) A piece of this tube, one metre long, is completely filled with lead. Find the value of the ratio

$$\frac{\text{volume of copper}}{\text{volume of lead}}$$
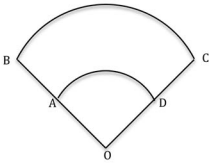
for this piece of tube.

(v) Given that the ratio of the masses of equal volumes of copper and lead is 4: 5, **write down** the value of the ratio

$$\frac{\text{mass of copper}}{\text{mass of lead}}$$

for this piece of tube.

Figure 2. Oxford and Cambridge Schools Examination Board, GCE Ordinary Level Mathematics Syllabus B, paper 2, Summer 1980, question 6.

A piece of material ABCD is to be used for covering a lampshade. OBC is a sector of a circle. Centre O, radius 30 cm. OAD is a sector of a circle, centre O, radius 12 cm. Angle BOC = 120°.

(c) Calculate the length of arc BC.

Arc BC = _____ cm

(3)

(d) Calculate the radius of the base circle of the lampshade which has the length BC for its circumference.

In this question give all your answers correct to 3 significant figures.

(a) Calculate the area of a circle radius 30 cm

Area = _____ cm^2

(3)

(b) Calculate the area of the shape ABCD.

Area = _____ cm^2

(4)

Radius = _____ cm

(3)

Figure 3. London East Anglian Group GCSE, Mathematics A, June 1991, Paper 4, question 10.
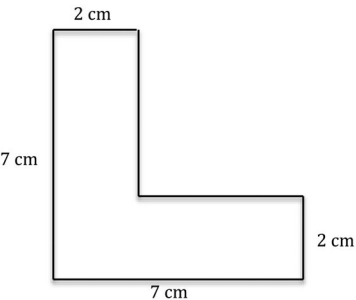
2 cm

Diagram NOT
accurately drawn

7 cm

2 cm

7 cm

The diagram shows the cross-section of a solid prism.
The length of the prism is 2 m.

The prism is made from metal.
The density of the metal is 8 grams per cm².

Work out the mass of the prism.

....
(Total 5 marks)

Figure 4. Edexcel, GCSE Mathematics (Linear) – 1380, Paper 3, Higher Tier, June 2011, question 10.

for the examinee's independent decision-making. Choosing characteristics that should be included in comparative analysis is the basic challenge for those trying to design an analytic scheme.

While designing our methodology, we also needed to find a way of translating such one-to-one contrasts as those instantiated above into overall differences between examinations. Task 3, therefore, was to answer the question of *how the resulting analytic scheme should actually be applied to the entire corpus of data*.

Before describing the ways in which we implemented our three tasks: (1) choosing a sample; (2) devising a sensitive, fully operational method of analysis; and (3) devising the way of applying the analysis to entire examination papers), we devote the next section to conceptual issues that needed to be resolved first.

1.3. Preparing the conceptual framework: what requirements must it fulfil?

Research methodologies are not standalone constructs; rather, they are derivatives of the way researchers conceptualise their object of inquiry, such as learning, teaching, or mathematics. Research methods are both enabled and constrained by the researcher's language. In the EDSM project, we derived the requirements with regard to a conceptual framework from our critique of the tacit assumptions that guided our own common-sense attempt at analysing the examination questions presented in Figures 1, 2, 3, and 4. These assumptions, we believe, are also widely present in current research in mathematics education. In section 1.2 our initial claim was that the four examination questions differed in their *mathematical content* and *level of difficulty*. Later we added more detailed features, such as the *guidance and support* provided for students, *the complexity* of language, the use of *diagrams*, and in the way *non-mathematical context* was implicated in each one of them. Below, we present three basic weaknesses of this kind of analysis and point to those properties of the underlying conceptual system that can be held responsible for these shortcomings.

The first weakness to note is that the above analysis may not be *subtle* enough to allow us to pinpoint those dissimilarities between examination questions that can make a difference to the student's vision of mathematics and to their performance. As has been shown in previous discursive research, non-identical statements regarded by a mathematically versed person as having 'the same mathematical content' may be seen by the student as anything but equivalent (see e.g. Sfard & Lavie, 2005; cf. Bezemer & Kress, 2008). This means that the analyst who uses such 'objectifying' words as *content*, *knowledge* or *concept* may be ignoring differences that she deems as 'content-preserving' yet can nevertheless alter the examinee's performance.

The content-form duality is so deeply rooted in our professional and colloquial languages that it may not be easy to remove. The majority of approaches widely applied in mathematics education research adopt it uncritically. The assumption that one can distinguish between *what* it is being taught and *how* it is taught, between 'mathematics-as-such' and the way it is 'presented', and between 'knowledge' and its 'representations' is obvious, for instance, in those studies that purport to be testing new, possibly better, *ways to teach* a given *mathematical content* or *concept* (see, for instance, "process-product" research [Kilpatrick, 2015]). Here, the researcher seems to be acting on the basis of the premise that one can change the form – the way of teaching or testing – while keeping the *content* – "the mathematics" – intact. The same assumption seems to underlie those studies that, while attending to the language of the examinations, have suggested that the *difficulty* can vary while the content that is being tested stays the same. Such research has produced a list of textual factors that, according to the authors, are likely to affect the difficulty of questions. These include, for example, the structure of the question (Pollitt, Hughes, Ahmed, Fisher-Hoch, & Bramley, 1998), use of diagrams, technical notation and language, the number of steps required and the demand for recall of knowledge or strategies (Fisher-Hoch, Hughes, & Bramley, 1997). These and other similar studies have influenced the practice of designing examination papers, in which the designers have sought to ensure that the language does not 'get in the way of the mathematics'.

In the light of all this, *our first requirement with regard to the conceptual approach to guide a study such as ours is that it be non-dualistic, that is, free from the unhelpful content/form dichotomy.*

The second weakness of our initial analysis, tightly related to the first, is that our comparisons were not grounded in clear, publically accessible criteria. The researcher's language, once again, seems to be the main culprit. The characteristics being compared are not operational, either inherently so or because they have not been properly defined. The term *content* seems to belong to the former category, whereas *complexity of language* is an example of the latter. Further, although properties such as *complexity* or *difficulty* sound quantifiable, they have not been accorded any numerical measures, and thus remain too vague to allow reliable, defendable comparisons. Users of such undefinable or undefined terms rely on the unwarranted assumption that the author and addressee share a common experience of school mathematics and hold similar views on how features such as 'complexity of language' or 'difficulty' may vary. In fact, claims made in these terms admit of many different, possibly contradictory interpretations.

Our second requirement, therefore, with regard to a conceptual framework for our kind of study is that the descriptors of examinations used in analyses be operationalised by being defined with the help of specific textual indicators, *such as linguistic components and structures of the text, well-delineated properties of diagrams, spatial organisation, etc.*

A proper analysis must differ from what was exemplified in section 1.2 in yet another respect: the units of text that are being compared. In the illustrative analysis above, we contrasted questions that we considered dealt with the same curricular topic and were somewhat similar to each other in the nature of the required activity. It is tempting to believe that such questions can count as tools for assessing the same aspect of the examinee's mathematics. In fact, findings that emerge from a question-to-question comparison cannot be seen as sufficiently representative. To claim otherwise is to expose oneself to the accusation of 'cherry picking' – of mistaking a non-randomly sampled local phenomenon for a general trend. Overall comparison, to be sound, would have to include quantitative analyses of the entire corpus.

The last requirement for our conceptual framework, therefore, is that it allows for translating characteristics of questions into features of the examination as a whole.

In the next part of this article we present the discursive approach that we have chosen as a conceptual framework for our research. We do this by implementing the three tasks listed in section 1.2 in accord with the three requirements discussed above.

2. Theoretical foundations: school mathematics as a discourse

In choosing a theoretical perspective for our project we turned to two conceptual frameworks that seemed to fulfil the condition of non-duality and have already been in use in research in mathematics education. The first is based in social semiotics (Halliday, 1978, 2003; Hodge & Kress, 1988). This approach has been used in science education research (Lemke, 1993) and in research in mathematics education (e.g. Herbel-Eisenmann & Wagner, 2007; Morgan, 2006, 2009; O'Halloran, 2005). The other is Sfard's (2008) communicational theory, which belongs to sociocultural tradition. Inspired by the work of Vygotsky and by Wittgenstein's late philosophy, it views language-based communication as central to all human activities. It has been used mainly in research in mathematics

education (see for example, two special issues: Sfard, 2012; Nachlieli and Tabach, 2016), but has lately started to appear also in science education (Rap & Blonder, 2016). Although these two approaches are rooted in different traditions, they have a great deal in common. In what follows we outline each separately and then describe how we combined them to create a framework for our study of change in GCSE examinations.

2.1 Social Semiotics

An important foundational tenet of social semiotics, one that allows it to avoid the content-form dichotomy, is that language and other communicational modes are functional, not representational. This entails that analysis of instances of communication (texts) does not attempt to uncover the intentions of the author (speaker/writer) or to determine some absolute 'real meaning' of the words or an aspect of the 'real world' that they refer to; rather, it focuses on what is achieved by the text within a particular context – in the case of the examinations we consider here, in the contexts of preparing for and sitting examinations. In general, communicational acts perform three types of function (named *metafunctions*): *ideational*, construing the 'reality' of the world; *interpersonal*, construing the identities and relationships of the participants in the communication; and *textual*, construing the role of the text itself as part of a social practice (Halliday, 1978).

With regard to the ideational function, Halliday and Mathiessen (1999) propose a definition of experience as "the reality we construe for ourselves by means of language" (p. 3), thus rejecting the traditional cognitivist perspective that language is a (more-or-less imperfect) means of representing pre-existing conceptual structures. This perspective thus orients us to study use of language as a means of understanding the ways in which the participants in communication may construe their 'reality'. As our concern is with mathematics and mathematics education, our primary interest in this ideational metafunction of language lies in the 'reality' of mathematics: What kinds of objects are dealt with in mathematical practices? What kinds of activities count as mathematical? Who (or what objects) are the agents in mathematical actions? What forms of reasoning are used? Where do mathematical facts come from?

Studying use of language with the focus on its interpersonal function allows us insight into the question of how participants position themselves and others within a social practice. Within our project, our main concern with respect to the interpersonal metafunction was to consider how the examinations position students with respect to mathematics and towards other participants of mathematical discourse: What kinds of mathematical activities are students expected (or not expected) to engage in? How autonomous are they in performing these activities? Where does authority lie (with the examiner, the student, the logic of mathematics)? These interpersonal questions have some similarity to those posed by Herbel-Eisenmann and Wagner (2007) from a similar theoretical perspective. See also the fuller account of a social semiotic approach to mathematics education in Morgan (2006).

Social semiotics offers not only a theoretical perspective on the functions of communication in general but also, with systemic functional linguistics (SFL), it provides a toolkit for analysing the functioning of specific texts. In the case of verbal language, it provides a means of relating the lexicogrammatical aspects of the text to each of the metafunctions of

language (Halliday, 1985). In recent years, there has been increasing attention to communication that uses multiple modes in addition to or in place of language, especially visual modes (e.g. Kress & van Leeuwen, 2001). Mathematical communication involves significant use of non-verbal modes, with important roles played by algebraic notation, diagrams, graphs and other specialised semiotic systems. Although there are some recent developments of social semiotic approaches for analysing the functions of various non-verbal modes used in mathematics (Alshwaikh, 2011; O'Halloran, 2005), we have not fully addressed the multimodal nature of examination texts in developing our analytic scheme. The article by Alshwaikh in this issue illustrates how multimodality may be incorporated more fully.

2.2. Communicational approach

The communicational framework,[1] as presented in Sfard (2008), is rooted in the claim that *mathematics may be usefully conceptualised as a discourse* and that *mathematical thinking is a form of communicating*.

Within this approach, the term *discourse* is to be understood as referring to a form of communication made distinctive by four characteristics: vocabulary and syntax, visual mediators, routines and endorsed narratives. Number words, names of arithmetic operations and names of geometric shapes are commonly recognised as typical parts of *mathematical vocabulary*. Although these words appear also in colloquial discourses, their use in mathematics, usually governed by explicit definitions, is often quite different. Mathematical *visual mediators* are physical objects with the help of which the participants of mathematical discourse try to make clear what they are talking about. Unlike some visual mediators that have come into being and exist in the world independently of discourse, mathematical mediators have been created specifically for the sake of communication. These include different kinds of symbolic artefacts, some of which have been mentioned above as constituting specialised semiotic systems: written words, algebraic ideographs, diagrams, graphs, and various iconic drawings. *Mathematical routines*, that is, patterned ways of performing mathematical tasks, may be described either by means of algorithms that determine the performance in a unique way (think, for instance, about calculating sums or multiples of integers) or by sets of rules that merely constrain the performer's actions (e.g. for proving theorems or defining new mathematical terms). Finally, the term *endorsed narratives* refers to any 'story' considered by a mathematical community as a useful and reliable description of what this community regards as the 'mathematical universe', populated by 'mathematical objects'. Narratives already endorsed become a basis for constructing more such narratives.

The term *mathematical object* used above is interpreted within this framework as referring to special discursive constructs created by means of metaphorical projection from discourses on physical reality. In addition to utterances about mathematical objects such as numbers, sets or functions, classroom mathematical discourse contains also utterances about people. This sets it apart from the formal mathematical discourse practiced in academia, which tends to exclude the human factor (though see Burton and Morgan's (2000) analysis of mathematics research papers which shows examples of exceptions to this tendency). In the texts of mathematical examinations, the presence of the human actor is conspicuous. Along with 'mathematising', that is, telling stories about mathematical objects,

examination texts tend to 'subjectify', that is, narrate people and their actions. Sentences such as "[I]t is proposed to reduce the disc to the required weight of 11 lb." (Figure 1) or "Write down an expression" (Figure 2) are good examples of subjectifying sentences (note that the first sentence is about human action, that of proposing, even though it does not mention any person explicitly). A single utterance may belong to both these categories – think, for instance, about sentences such as "When I add 2 and 3, I get 5". In Halliday's terms, such utterances perform both ideational and interpersonal functions. However, although the terms in each of the two pairs, ideational/mathematising and interpersonal/subjectifying are related, they are not the same. Mathematising and subjectifying refer to *kinds of discourse* (about mathematical objects and about people, respectively), whereas Halliday's distinction regards *functions* that every text fulfils, in one way or another. Still, for many purposes these two pairs of terms are interchangeable.

2.3. Discursive framework for the study of changes in GCSE examinations

The last sentences of the previous section alert us to a 'family resemblance' between social semiotics and the communicational approach. Both these frameworks focus on communication as the activity which, being present in, and central to, all human processes, may well be the primary source of all things human. Researchers from these two schools are interested in the same types of phenomena and ask similar questions. Further, as already stated, both approaches agree in their rejection of the content-form dichotomy, even if only the communicational approach argues for this rejection in an explicit way. Finally, while each of the two approaches offers its own set of conceptual and methodological tools, these sets are fully compatible and sometimes even exchangeable.

While deciding about the conceptual framework for the EDSM project we considered the possibility of using just one of these approaches. We soon realised that giving up either might be a waste, because the two frameworks are complementary: each has something useful that the other is missing. In the most general terms, the conceptual and methodological tools of the communicational framework are tailor-made for the study of mathematical discourse and, as such, were a natural choice for analysis of changes in mathematics examinations. Social semiotics, on the other hand, has a rich set of generic tools for dealing with the subtlest aspects of discourses, and we thus expected to be able to refine our analysis with their help.

Keeping all this in mind, we opted for taking advantage of both frameworks. Yet, rather than presenting the combined framework explicitly, we decided to build it gradually, as we go; we looked at the social semiotic and communicational approaches as resources, each of which could be used at will at any stage of our study, according to need (and, inevitably, according to personal preferences). Because of the ontological and epistemological commensurability of the two frameworks, we did not fear that this might lead to inconsistencies. We had, however, to remain wary of the risk of double terminology. In the rest of this article we refer to our emerging combined framework as *discursive*.

The discursive framework can be shown to fulfil the three requirements formulated in section 2.3. Its non-duality has already been noted. Let us remark now that this feature makes the researcher able to deal with changes that are unlikely to be noticed without the discursive lens, yet are sufficiently influential to be worth attention. Indeed, discursive investigators, reluctant to unify different texts under the title of 'the same content', do not

dismiss any change in wording, syntax or structure of examination questions before asking themselves whether, and under what circumstances, this difference may have an impact on students' interpretation or response. Consequently, a discursive approach differs from other, more traditional ones in the way it divides observed phenomena into 'the same' or 'different'. Not only are the divisions subtler, they also run along different lines.

The second requirement, that of *operationality*, is fulfilled as well. By equating mathematics with a form of communication, the discursive approach dispenses with problematic dichotomies and their nebulous ingredients. Since our object of study is the discourse and the texts it produces, its descriptors can be defined with the help of specific textual indicators. By freeing us in this way from assumptions about "common understanding" of texts, this approach protects us from being misled by our own spontaneous interpretations. As researchers, we are now now able to interpret the text from both the position of an *insider* and *outsider* to our own mathematical discourse (Fairclough, 2001; Morgan, 2014b; Sfard, 2008, 2013).

Once analysis of GCSE examinations is seen as a task of characterising texts according to well-defined, publicly accessible indicators, it becomes clear that the last requirement, *translatability of properties of questions into features of the examination as a whole*, is fulfilled as well. Working with each indicator at a time, we can now attach codes to different units of text: individual words, phrases, sentences, sub-questions/tasks or complete questions (units of different size are appropriate for different indicators). The coding process will result in the production of a database that can be investigated for discursive properties of each examination as a whole, and then for variation in these properties across examinations. This completes the justification of our choice of conceptualisation to guide the design of our research method. The description of the way we actually built this method comes next.

3. Developing the method

In this section, we introduce our analytic method while explaining how we implemented the three tasks presented in the conclusion of section 1.2.

3.1. Constructing the database

For our project, we needed a data corpus that was manageable in quantity, but also extensive enough to allow us arrive at a reliable picture of how the discourse of GCSE examinations has changed over a period of curriculum reform. We had to decide the time interval to study and which examination papers and items to include from within the time interval. The examination system in England is diverse in that there are several examination boards that produce 'equivalent' but competing examinations. It also provides different examination papers for students perceived to have different levels of attainment. Further, the structure of the system has changed over time, partly, but not entirely, in parallel with curriculum changes. All these factors had to be taken into account in constructing our database.

The first decision to make regarded the time interval for the study. Our initial question was *How has the discourse of the GCSE examination changed since its inception?* However, the introduction of the GCSE was itself one of a series of reforms following the publication

of the Cockcroft Report (DES, 1982). Recognising the major impact of this report on mathematics education reform in England, we decided to take 1980, two years before its publication and eight years before the first GCSE examinations, as the starting point for our study. A review was then undertaken of notable changes in the curriculum, policy, in examination specifications and in public debates about examinations since 1980. This enabled us to choose dates most relevant to our task of mapping the evolution in GCSE examinations. This resulted in the selection of years shown in Table 1.

The examinations in England prior to 1988 consisted of two separate systems, the GCE O Level taken by higher attaining students and the CSE taken by others. The GCSE was designed as a single qualification system but students at different levels of attainment take different examination papers. In the EDSM project, limited resources led us to decide to focus only on the examinations for higher attaining students. This choice of focus necessarily limits the scope of the conclusions of the study.

For each chosen year we included the complete examination papers set at the summer sitting from the most popular syllabuses of two of the three main examination boards in England. In order to be able to interpret what the examinations expected from students as producers as well as consumers of mathematical discourse, we also collected mark schemes and examiners' reports where these were available for the chosen examinations.

3.2. Developing an analytic scheme

Each of the examination papers in the resulting database had now to be described according to a well-defined set of criteria which, when taken together, would constitute our *analytic scheme*. To develop the scheme, we needed to: (1) specify *the aspects of discourse* on which the analysis should focus; (2) formulate *questions* about each of these aspects; and (3) operationalise these questions by defining their central notions with the help of *textual indicators*.

Designing an analytic scheme is an iterative process of gradual refinement, in which each act of asking a question is followed by an attempt at operationalisation and each such attempt is then likely to lead to a revision of the question. In our study, we began by inspecting examples of examination questions in order to develop an informal sense

Table 1. Years of examinations included in the EDSM data base.

Year	Reasons for choosing this year
1980	As indicated above, we chose this year as a pre-Cockcroft Report base-line GCE 'O' level
1987	This was the final examination prior to GCSE. We chose a transitional syllabus in this year: a 'joint' GCE/CSE examination.
1991	This was the third year in which GCSE examinations were set. Students taking the examination this year were the first cohort to have been prepared for GCSE through the whole of their secondary school education.
1995	Students taking the examination this year were the second cohort to have followed the National Curriculum throughout their secondary education. A report by the School Curriculum and Assessment Authority following the examination this year called for increased emphasis on algebra and non-calculator question papers (SCAA, 1996).
1999	This year was chosen to reflect changes made to examinations as a result of the SCAA (1996) recommendations.
2004	A report by the Qualification and Curriculum Authority remarked that question papers this year were 'more accessible linguistically' than in 1999 and 'more clearly laid out' (QCA, 2006, p. 10).
2010	The final year of examinations before the start of the project.
2011	Added to the data base after the start of the project. Informal discussion with an ex-official of the QCA suggested that examination boards were introducing less structured questions in 2010 and 2011.

of differences between them. The resulting initial characterisation was informed by knowledge of characteristics of mathematical discourse as identified in existing literature in linguistics and in mathematics education. This led to drafting an initial set of descriptors, which were then applied to further examples. We scrutinised the results to see if we managed to capture distinctive discursive features of each examination question. If not, the whole procedure was repeated, resulting in a refined, although perhaps still not fully satisfactory, set of descriptors.

Early on in this meandering process we decided on two sets of foci. Mindful of the fact that any utterance contributes to stories about both mathematical objects and human participants in mathematical discourse, we resolved to attend to these two types of storytelling separately. In one part of the analysis, we would describe the examination texts according to how they do the work of *mathematising*: We would inquire about what *stories about mathematical objects* are told, either explicitly or implicitly, in examination questions, and about how they are told. This part of analysis may be described in colloquial terms as aiming at a description of 'the mathematics' involved in the examinations. The other part of the analysis would focus on *subjectifying* and, more specifically, on what can be told on the basis of the examination questions about students and their expected participation in mathematical discourse: how they are to engage with the text and how their subsequent problem solving activity is supposed to proceed.[2]

For each of the two parts of our future analytic scheme we now needed to: (I) list the aspects of the discourse we deemed as worth considering; and then (II) ask questions about each of these aspects. To help readers navigate through the following description of these two layers, we introduce Tables 2 and 3 that summarise, respectively, the mathematising and subjectifying parts of the resulting analytic scheme. Layers I and II can be seen in the first two columns of each table. Column III operationalises each question by specifying the relevant textual indicators.

To be able to characterise the work of mathematising done in the text of examination questions (Table 2), we decided to focus on the four basic features that make discourse mathematical: the use of mathematical words and of visual mediation, mathematical routines, and narratives about mathematical objects. While speaking about the use of *words* (Table 2(IA)), we were interested not so much in knowing what the specialised mathematical words were as in the question of the *extent to which specialised language was used* (IIA). We also wished to assess whether, and to what degree, the discourse was *objectified* (IB): we asked whether mathematical stories told by the examiner (and those yet to be told by the examinee) were predominantly about properties of independently existing mathematical *objects* or about *processes* that take place over time. Almost any mathematical statement about objects can be translated into an equivalent statement about processes (Sfard, 1991). For instance, the claim "5 is the limit of (5x+1)/x in infinity" that speaks about the object called *limit* is equivalent to "(5x+1)/x tends to 5 when x tends to infinity", in which this object does not appear. The question of the degree of objectification is of importance, because objectification is often a condition for a further development of mathematical discourse (the growth of this discourse is the iterative process of turning processes into objects, and then studying processes on these new objects [Sfard, 1991, 2008]).

The next illustration explains the section of Table 2 pertaining to endorsed narratives. These narratives are *stories about mathematical objects* told in the examinations (Table 2

Table 2. Analytic framework for mathematising aspects of examination discourse.

I. Aspects of the discourse	II. Questions guiding the analysis	III. Textual indicators
Vocabulary and syntax (lexico-grammatical aspects)		
A. specialisation	To what extent is specialised mathematical language used?	• lexical items used in accordance with mathematical definitions, considered at the level of: ○ vocabulary ○ sentence ○ text unit • extra-mathematical context ○ depth of engagement with context
B. objectification of the discourse	To what extent does the discourse speak of properties of objects and relations between them rather than of processes?	• nominalisation: use of a 'grammatical metaphor', converting a process (verb, e.g. *rotate*) into an object (noun, e.g. *rotation*) • the use of specialised mathematical nouns such as *function*, *sequence* which encapsulate processes into an object • complexity of compound nominal groups
C. logical complexity	What kinds of logical relationships are present and how explicit are they?	• the types and frequencies of conjunctions, disjunctions, implications, negations and quantifiers
Visual mediators		
D. the presence of multiple visual mediators	To what extent does the discourse make use of specialised mathematical modes?	• presence of tables, diagrams, algebraic notation, etc.
	How are multiple visual mediators incorporated into the discourse?	• provided in the text or to be produced by the student • linguistic, visual and/or spatial relationships between modes
E. transitions between visual mediators	What transformations need to be made between different modes?	• presence of or demand for two or more modes of communicating 'equivalent' information, e.g. an equation formed from a word problem; a unit of text that involves table, graph and algebraic expressions corresponding to the same function
	How are transformations indicated in the discourse?	• provided in the text or to be produced by the student • explicit linguistic or visual links between modes
Routines		
F. the types of action demanded of students	What areas of mathematics are involved?	• topics
	What are the characteristics of the routine procedures?	• algorithmic or heuristic? • complexity • explicitly hinted at?

(*Continued*)

Table 2. Continued.

I. Aspects of the discourse	II. Questions guiding the analysis	III. Textual indicators
Endorsed narratives		

G. the origin of mathematical knowledge	What is the degree of alienation of the discourse?	• mathematical objects as agents in processes
		• agency obscured by:
		  o non-finite verb forms
		  o passive voice

	To what extent is mathematics construed as involving material action or as atemporal objects and their properties?	• mathematical objects involved in:
		  o material processes
		  o relational or existential processes

	To what extent is mathematics presented as a human activity?	• human agents in mathematical processes
		  o thinking
		  o scribbling

H. the status of mathematical knowledge as absolute or contingent	To what extent does the text indicate that decisions or choices are possible during mathematical activity?	• modifiers indicating degree of certainty (e. g. may, can, will …) • conditional clauses (e.g. if … or when …) • explicit decisions have been or need to be made

(G) and 2(H)). The most interesting of the stories are those that regard the onto-epistemological status of these abstract objects, that is, tell us about their origins, their relation to the world and the ways to construct and endorse new claims about them. These foundational stories, rather than being told explicitly, can be derived from the way mathematical objects are talked about. It is this way of talking that implies what answers can, or cannot, be reasonably given to questions such as "To what extent does mathematics involve material action or atemporal objects and their properties?" (Table 2(IIG)). Statements about relational or existential processes, such as "3 plus 4 gives 7" or "There exists a continuous nowhere differentiable function" involve no human actions. Indeed, these sentences are fully *alienated*, that is, free from human presence, and feature mathematical objects as independent agents. On the other hand, utterances such as "If we add 3 to 4, we obtain 7" or "It is possible to construct a continuous function that is not differentiable at any point" are more consistent with the claim that operating on mathematical objects is an evolving human activity. The way the foundational story of mathematics is deduced from examination questions is elaborated and instantiated by Morgan (this issue).

To characterise the work of subjectifying done by examination texts we focus on two aspects of the interaction: on the relationship between the student and the examination author (Table 3(A)) and on the degree of the student's autonomy in mathematical discourse, that is, on the question of how free she is to make her own problem-solving decisions (Table 3(B)).

The way we came across this last issue is worth telling. At an early exploratory stage we noticed that towards the extremes of our timeframe (1980 to 2011), examinations

Table 3. Analytic scheme for subjectifying aspects of examination discourse.

I. Aspects of the discourse	II. Questions guiding the analysis	III. Textual indicators

A. student–author relationship	What kind of relationship is constructed between the student and a mathematical community?	• use of personal pronouns
		  ○ inclusive or exclusive we
		  ○ other personal pronouns

	Is the student given instructions or invited to consider mathematical questions?	• interrogative (questions) • imperative (instructions)
B. student autonomy	In responding to an examination question, how many independent decisions is the student allowed/required to make in:	• the grain size of the task
	• designing the path to follow?	

	• interpreting the task?	• complexity of utterances
		  ○ lengths of a sentence
		  ○ grammatical complexity: the depth of 'nesting' of subordinate clauses and phrases
		  ○ logical complexity

	• choosing the form of the 'answer'?	• the layout
		  ○ the physical size of the answer
		  ○ the space provided for the work to be done on the way toward solution
		  ○ format of the answer (units, precision, no. of solutions)
		  ○ modality of the answer (graph? algebraic expression?)

	• choosing/constructing the mode of response?	• visual mediators: verbal, symbolic, or graphic: supplied or to be produced?

appeared to vary in the nature and amount of guidance provided to structure the student's response. In the later examinations, there appeared a clear limitation on the examinee's freedom in deciding about the problem-solving trajectory and on the format of expected answers. This decrease in students' independence meant a consequential change in the nature of activity they were expected to perform on the basis of their own decisions. These observations made it clear that, in our analysis, we need to find out how the labour of solving mathematical problems is divided between the examiner and the examinee and what kinds of mathematical activities the students are expected to be able to perform independently (this is elaborated further in example 2 in section 3.3).

As already explained, in developing the indicators presented in column III of Tables 2 and 3, we tried to formalise our intuitions rather than abandon them altogether. We also made use of tools that had previously been developed as descriptors of mathematical discourse within the communicational approach (Sfard, 2008), combined with components of SFL as applied to mathematical texts by Morgan (1998, 2006).

In considering the way indicators help in operationalising the analysis-guiding questions, note that some of the properties of discourse referred to in the questions posed in column II of both Tables 2 and 3 are purely qualitative, and as such, their answers may be read directly from the text (see for instance the question "What areas of mathematics are involved?", Table 2(F)). Other properties, such as logical complexity (Table 2(C)) or the degree of objectification (Table 2(B)), are quantitative and would have to be calculated from what is visible in the text with the help of explicitly defined procedures. For instance, in the article by Morgan and Tang in this issue the authors explain how the degree of objectification can be quantified. In the next section of this article we elaborate and instantiate some of the methods of assessing the aspect *student autonomy* (Table 3(B)) through quantifying the decision making required in interpreting a task (using the indicator *grammatical complexity*) and in designing a solution path (using the indicator *grain size of the task*).

At this point it is important to remember that our analytic scheme was designed to provide a well-defined, reliable set of tools that could be used not only to produce detailed qualitative analysis of individual questions, but also to code a large sample of examination papers. Thus, for instance, to find out how examinees' autonomy has been changing over years, we would now look at every examination in its entirety, that is, translate characteristics of separate questions, such as complexity of utterances or specific properties of the layout, into a feature of the examination as a whole. As will be exemplified in section 3.3, this can be done in several ways.

Of course, quantitative analysis by itself does not tell the whole story. Each individual examination question involves a set of multiple discursive characteristics. If we are interested in how students' experience of mathematics may have changed over time – if we wish to follow the evolution of the stories about mathematics and about themselves that students were likely to 'read' from the examinations – it is relevant not only to investigate how single characteristics may have varied, but also how the characteristics were combined into complete texts. This requires a qualitative approach that allows the analyst to describe individual questions with the help of the range of properties in the analytic scheme. Such descriptions then give rise to a richer and more nuanced interpretation of expected student engagement with mathematics and mathematical activity. Qualitative analyses of individual questions helped us in the later phases of the project, where we tried to find out how different discursive characteristics affect student participation in mathematical discourse. Based on those analyses, we chose and designed pairs of examination questions representing the 'old' and 'new' types of GCSE examinations and administered them to students in written tests and task-based interviews (see Morgan, 2014a; Morgan, Tang, & Sfard, 2012 for preliminary analysis of some of the outcomes of this part of the project).

3.3. Applying the scheme

In this section, we complete the presentation of our method with two examples showing how the analytic scheme presented above was applied to our data.

Example 1: Grammatical complexity

Earlier in this article we claimed that the four questions presented in Figures 1, 2, 3, and 4 varied in the complexity of their language. Such complexity contributes to the formation

of mathematical objects (objectification of the discourse) and thus appears in the mathematising part of our analytic scheme (Table 2(B)). It appears again in the analysis of subjectifying (Table 3(B)) because it is related to the number of decisions the examinee has to make while interpreting the text of the examination and is thus an indicator of student autonomy.

Two types of complexity are included in the part of the scheme shown in B in Table 3, that of the grammatical structure of the text and that of the logical connections that are visible through the use of logical connectives such as 'and', 'or', 'if … then', because', etc. As a measure of this latter feature we used *the average number of connectives per sentence*. This average was calculated for each year in our sample. As shown in Table 4, the thus measured logical complexity changed rather dramatically between 1980 and 2011: at the end of that period, the average number of logical connectives in a sentence was half what it was at the beginning.

The other type of complexity, that of the grammatical structure of the text, can be measured in a number of ways. Probably the most immediate indicator, and also the easiest to find, is *the length of the sentence*. Our initial informal explorations left us with the impression that as time went by, the length of sentences in examination papers went down. We corroborate this visually gleaned difference with the help of a more formal analysis of questions from examinations set in 1980 and 2007,[3] presented in Figures 5 and 6.

Indeed, the instructions in the 2007 question are composed either of a single word or of a pair of words ("expand", "factorise", "solve completely"), whereas in the 1980 question, they are between two and 11 words long. The average numbers of words per instruction are 6.3 in 1980 and 1.2 in 2007. Table 5 shows the difference between two examination papers from 1987 and two from 2011, displaying not only the average, but also the maximal length of a sentence in an examination paper. Again, these measures of complexity are lower for the later examinations.

Another, more sophisticated, but also more informative indicator of grammatical complexity is one that draws on the property of recursivity of language, that is, on the fact that we often build our sentences by replacing a word with a compound phrase. Thus, we may say "ET did something", but we may also say "The alien who came from another planet did what it was asked to do", with the latter sentence obtained from the former by substituting "The alien who came from another planet" instead of ET and "what it was asked to do" instead of "something". The procedure is recursive, in that it would be possible to complicate the latter sentence even further by replacing different words with compound expressions (for instance, instead of "planet" we could write "heavenly body not unlike

Table 4. Average number of logical connectives per sentence as a function of time.

year	1980	1987	1991	1995	1999	2004	2010	2011
Average number of logical connectives per sentence	0.68	0.48	0.42	0.39	0.37	0.33	0.35	0.34

Notes: The appearance of *given* or *given that* used as a connective has been calculated separately and it is not included in the evaluation of logical complexity shown in this table. These terms were not identified as relevant to this analysis until a late stage in the project. *Given*, used as a connective, appeared extensively in the examination papers from 1980 (23 instances) and 1991 (11 instances), but was found only in three instances in 2011. Its almost complete eradication is consistent with the more general simplification of the complexity of the syntax of recent examination papers evidenced in this table.

(a) Factorise completely $ab^2c^3 + bc^2a^3$.

(b) Given that $r = \dfrac{s}{r+s}$, express s in terms of r.

(c) Simplify $\dfrac{v}{v-3} + \dfrac{v}{v+3} + \dfrac{18}{v^2-9}$, giving your answer as a single fraction in its lowest terms

Figure 5. Oxford and Cambridge Schools Examination Board, GCE Ordinary Level Mathematics Syllabus B, Summer 1980, paper 1, question 2.

(a) Expand $p(p^2 - 3p)$

----

(2)

(b) Factorise $y^2 + 5y$

----

(2)

(c) Factorise completely $2x^2 + 6xy$

----

(2)

(d) Solve $x^2 - 2x - 15 = 0$

----

(2)

(Total 8 marks)

Figure 6. Edexcel GCSE Mathematics Higher Tier Specimen Paper, 1987, paper 3, question 7.

Table 5. Average and maximum number of words in a sentence for two examination papers from 1987 and two from 2011.

Year/examination	1987a	1987b	2011a	2011b
Average number of words in a sentence	13.5	14.5	8.6	8.3
Number of words in the longest sentence	37	42	22	27

Earth"). One way to assess complexity of a sentence would thus be measuring the depth of such recursive nesting.

Another, related, method of assessing complexity to be considered here, is to look at the average length of the nominal groups, such as "The alien who came from a heavenly body not unlike Earth" or "the angle subtended by the chord AB at the centre of the circle" or "the sum of the first four even numbers". A compound nominal group is a phrase with more than one word that plays the same grammatical role as a single word naming an object or concept (a noun). Such a phrase can be the subject or object of a sentence, may be assigned properties, etc. Our interest in this particular indicator stems from the fact that the use of compound nominal groups to construct and name objects is typical of the language of mathematics and science. Nominal groups are grammatical devices for packing a large amount of information into a single grammatical unit, thus contributing to the cognitive demand of reading and interpretation. However, they also play an important role in shaping students' response and in creating potential for further mathematical activity. Above all, they invite the student to one of the most mathematical of mathematical activities: to the process of compressing the discourse in order to be able to say more with less. Extensive use of compound nominal groups thus contributes to the process of objectification of mathematical discourse and, more generally, to construing our experience of the world in terms of objects, their properties and relationships between them rather than in terms of actions and processes (Halliday, 1998; Sfard, 1991, 2008). At the same time, it introduces new mathematical objects that can themselves be assigned further properties and can act and be acted upon. As such, this contributes to the mathematising aspect of the discourse and is included in section B of Table 2.

We will demonstrate the use of this indicator with the help of the four examination questions we introduced at the beginning of this article in Figures 1 to 4 above. The longest nominal groups in each of the four questions are shown in Table 6. A simple

Table 6. Compound nominal groups.

	longest nominal groups	no. of words
Figure 1 (1957)	• a uniform circular metal disc of diameter 15 in	9
	• a number of ¼ in. diameter holes	7
Figure 2 (1980)	• the volume of copper in a piece of the tube one metre long	13
	• the ratio of the masses of equal volumes of copper and lead	12
Figure 3 (1991)	• a sector of a circle, centre O, radius 12 cm	10
	• the radius of the base circle of the lampshade which has the length BC for its diameter	17
Figure 4 (2011)	• the cross-section of a solid prism	7
	• the mass of the prism	5

count of the number of words in each nominal group reveals a marked difference between those occurring in 1980 and 1991 and those occurring in the questions from 1957 and 2011.

We cannot end this example without a disclaimer: the analyses we have used here to illustrate our methods suggest some differences in complexity between examinations in different years but we need to remember that the overall assessment of complexity will emerge by combining the results obtained with the help of different indicators and investigating patterns of variation across the full data set.

Example 2: Grain size of the task

Designing a problem-solving trajectory is probably the most obvious context for assessing student autonomy. We have defined the *grain size of the problem* (Table 3(B)) as the minimal number of decisions (choices) the problem solver must make while designing a series of elementary steps necessary to solve the problem. The descriptor *elementary* specifies a step that can be executed in a single automated operation, without further partitioning of the problem into smaller ones, so that no further decisions regarding its implementation are necessary. Of course, this definition is relative, since the answer to the question of whether a step can or cannot count as elementary depends on what may be defined as automated operation. For instance, squaring 12 may be an elementary step for some students (those who have memorised the result), but will be a compound move for those who need to perform a calculation. This relativity, however, should not worry us in the present context of *comparing* examination questions. For our purpose, it suffices that all the questions are compared according to the same criteria. In our evaluations, we will decide what is to be considered an elementary step according to our own current sense of what a 16-year-old student with a reasonable mastery of the mathematical discourse at hand is likely to be able to perform in a single step.

Comparing the questions from the 1980 and 2007 examinations shown, respectively, in Figures 5 and 6 will now be done solely for the sake of showing that two similar questions may differ considerably in the grain size of their solutions. The solutions of relevant sub-questions, showing our identification of elementary steps and hence of grain size, are presented in Table 7. Choosing these particular items is justified because they belong to the same curricular 'slot': both questions aim at testing the examinee's competence in solving some types of equations (5b and 6d) and in transforming algebraic expressions into equivalent ones by expanding them (6a), by factorising (5a, 6b, and 6c) or by simplifying (5c). The operations of factorising an expression and of solving an equation appear in both of them. Simplifying is found in the 1980 question, but not in the other one, whereas for expanding the situation is reversed. Even without any formal comparison it is clear that simplifying a long sum of algebraic fractions requires far more decisions than expanding the polynomial expression, and thus that the 'grain size' of the former operation is much greater than that of the latter one. In the more recent examination, therefore, the sub-question representing the smallest grain size has been added, replacing the one that exceeds in its complexity all the other sub-questions appearing in Table 7. This seems to signal that the 2007 question has a lower average grain size than its 1980 counterpart.

Comparisons between the corresponding factorising sub-questions, 5a versus 6b and 6c, and then between the two equations, 5b and 6d, show a similar relation. Table 7 lists the decisions the problem solver has to make in order to decompose the implementation into elementary steps (in the case of questions such as this one, which refers the

Table 7. Solutions of factorisation and equation tasks from examination questions presented in Figures 5 and 6.

Factorisation questions		Equations	
Figure 5 a $ab^2c^3 + bc^2a^3$	choose the highest n so that a^n can be used as a common factor	**Figure 5** b $r = s/(r + s)$	choose a transformation [operation on both sides: $x(r+s)$]
$= a(b^2c^3 + bc^2a^2)$	choose the highest n so that b^n can be used as a common factor	$r(r+s) = s$	choose transformation [expand]
$= ab(bc^3 + c^2a^2)$	choose the highest n so that c^n can be used as a common factor	$r^2 + rs = s$	choose transformation [operation on both sides: $-rs$]
$= abc^2(bc + a^2)$	3 *decisions*	$r^2 = s - rs$	choose transformation [factorise]
Figure 6 b $y^2 + 5y$	choose the highest n so that y^n can be used as a common factor	$r^2 = s(1-r)$	choose transformation [operation on both sides: divide by $1-r$]
$= y(y + 5)$ -	1 *decision*	$r^2/(1 - r) = s$	5 *decisions*
		(Remark: this problem should also include deciding on the numerical constraints imposed by the operation of dividing and concluding that if r=1, there is no solution for s; we doubt, however, whether the examiners expected this addition).	
Figure 6 c $2x^2 + 6xy$	choose the highest n that can be used as a common factor of 2 & 6	**Figure 6** d Option 1 Step 1: *decide* what needs to be substituted for a, b, and c in the formula $-b \pm \sqrt{b^2 - 4ac}/2a$; implement	
$= 2(x^2 + 3xy)$	choose the highest n so that x^n can be used as a common factor	Step 2: calculate the *numeric expression* obtained (No decision is needed: in algebraic problems such as this one, we are not interested in decisions involved in numerical calculations; thus, the only algebraic step here is the proper substitution of numerical values for the variables a, b, and c). 1 *decision* Option 2 Step 1: *choose* two new numbers Step 2: check whether the numbers you choose give -15 when multiplied and give 2 when added. If not, go to step 1. 1 *decision*	
$= 2x(x + 3y)$	2 *decisions*		

solver to an algorithm, the solver must choose one of several possibilities specified in advance by that algorithm). Both in the case of factorisation and of equation solving, the grain size of the 2007 sub-question is lower than that of the corresponding 1980 one(s).

We have analysed grain size of all tasks in the examinations for two of the years in our database (by *task* we understand a question or sub-question that is signalled as requiring an answer from the examinee by being labelled with a sub-letter or number, having an answer space or line to write on, having a number of marks allocated to it, etc.). The results for 1987 and 2011 shown in Table 8 indicate that the average grain-size of examination tasks shifted between these years toward the lower end of the scale: the percentage

Table 8. Distribution of tasks among the different grain sizes in examinations of 1987 and in 2011.

% of tasks with grain size:	1	2	3	4
1987	60.32	23.81	12.70	3.17
2011	60.83	27.66	6.38	2.13

Note: Given that since 2010 there has been pressure on examination boards to decrease structuring of questions, it seems likely that exams from a little earlier than 2011 (1990s and early 2000s) would show more marked lack of tasks with higher grain sizes.

of tasks requiring two steps grew, whereas the percentage of those requiring three or four steps decreased. This suggests that at least some of those tasks that the student was once expected to perform independently were in 2011 regarded as requiring scaffolding. Of course, to know whether this represents a trend over time, we would need to perform the analysis of grain size for more years.

4. Concluding remarks: studying change of discourse in examinations and beyond

In this article, drawing on the work of the EDSM project and on examples of examination questions taken from our data set, we have presented the theoretical and methodological foundations of a discursive approach to the study of school mathematics. We conclude this article by remarking on the applicability and feasibility of the resulting research method.

With regard to the first of these issues, that of applicability, one may wonder how general the method is. Our analytic scheme has been developed through interaction with the examination texts studied by the EDSM project and it may therefore be specific in some respects to this context. In particular, we recognise that the genre of examination question is distinctive and that school mathematics discourse as a whole includes texts in a wider range of genres that may have characteristics not captured by our analytical scheme. Moreover, our interpretation of how the student may read, interpret and respond to the examination texts is built upon our understanding of the context for which they are meant. Texts that belong to other school contexts are likely to involve different relationships between student, teacher and text. We are also fully aware that, being conducted in England, our study may have many characteristics that limit its immediate applicability in other countries. In spite of all this, we believe in the wider usefulness of our scheme. First, it may be applied, possibly in a slightly adapted form, in investigating other types of school mathematics texts. Second, it can be used both to study single educational contexts and to make comparisons between contexts: to investigate variations in texts over time, between national education systems, between educational provisions for students of different social backgrounds or perceived abilities, between the modes in which they appear (e.g. in written form in a textbook versus as spoken classroom discourse), perhaps even between school subjects. An example of adaptation and application of the scheme can be found in this special issue in the article by Jehad Alshwaikh. The importance of this tool and of analyses conducted with its help lies in the fact that the results may sensitise teachers, examination writers, curriculum developers and policymakers to the impact of textual factors that have so far escaped the attention of both researchers and practitioners. Finally, like the SFL-inspired framework proposed by Herbel-Eisenmann and Wagner (2007), our tool has potential to be used by teachers and students to engage critically with school mathematics texts and to consider how they might be different.[4]

As to the feasibility of our method, there is reason to wonder. While highly effective, this method is also extremely work-intensive and time-consuming, and we must ask whether insights gained with its help are worth the investment. This question is further justified by the fact that parts of our findings may seem an elaborate corroboration of what has been known for some time. Our response to this is that there is a difference between knowing *that* something is the case and knowing the exact nature of this something. Moreover, if one wishes to know *how* the change happens, our analyses yield precise

information about which aspects of mathematical discourse are affected, and allow us to interpret what these transformations mean in terms of our ability to attain what we consider as the goals of mathematics education. With the help of the scheme developed in the EDSM project, our analysis can be explicit about the nature of the changes that have been happening over three decades in school mathematics. Above all, we are able to identify in detail which aspects of mathematical discourse are being transformed, omitted or added. This analysis allows us to see whether vitally important qualities are likely to be absent in the mathematical toolbox with which the English education system equips those completing secondary education.

Moreover, thanks to the nature of our research tools, whenever we find a change that is potentially harmful, this finding comes together with means for repair. Equipped with a lens through which school mathematical discourse can be monitored in subtlest detail, teachers and examination designers may now be able to engage in a focused effort to develop those fine, but now well-defined aspects of mathematical discourse that have been neglected so far. Of course, whether such effort is going to be undertaken depends on many factors, with politics being among the most important of them (see the analysis of influential discourses in the article by Lerman and Adler in this issue). We hope that this project may, at least, provide evidence and tools to inform a more rigorous debate on the standards of school mathematics.

Notes

1. This framework is sometimes called commognitive, with this portmanteau signalling that any statement containing this term refers to both communication and cognition.
2. These two parts of our analytic scheme correspond, roughly, to the ideational and interpersonal functions of language. Our scheme does not pay separate attention to the third of Halliday's metafunctions, the textual function. Some features of the text that contribute to this function have been subsumed within the two major parts of the scheme. For example, the physical layout of the question, structuring students' engagement and answers, was included in the part focusing on the participants.
3. The 2007 example is taken from a sample paper for a new syllabus, published as guidance for teachers and students. It was not part of our main data set but was used during the development of the analytical scheme.
4. The scope of the EDSM tool is wider than that proposed by Herbel-Eisenmann and Wagner as it addresses mathematising aspects of the discourse as well as the ways that students are positioned within it.

Acknowledgements

Funding

References

Alshwaikh, J. (2011). *Geometrical diagrams as representation and communication: A functional analytic framework* (PhD thesis). Institute of Education, University of London, London.

Andrews, P., & Sayers, J. (2005). How do teachers of mathematics teach? A four-way international comparison. *Research in Mathematics Education*, 7(1), 139–158. doi:10.1080/14794800008520150

Barnes, M., Clarke, D., & Stephens, M. (2000). Assessment: The engine of systemic curricular reform? *Journal of Curriculum Studies*, 32(5), 623–650.

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication*, 25(2), 166–195. doi:10.1177/0741088307313177

Britton, E. D., & Raizen, S. A. (Eds.). (1996). *Examining the examinations: An international comparison of science and mathematics examinations for college bound students*. Dordrecht: Kluwer Academic Publishers.

Broadfoot, P. M. (1996). *Education, assessment and society*. Buckingham: Open University Press.

Burton, L., & Morgan, C. (2000). Mathematicians writing. *Journal for Research in Mathematics Education*, 31(4), 429–453.

Clarke, D., Keitel, C., & Shimizu, Y. (Eds.). (2006). *Mathematics classrooms in twelve countries: The insider's perspective*. Rotterdam: Sense Publishers.

Department for Education. (2014). Higher standards as thousands more take GCSEs at the right time. *Press release*. Retrieved from https://www.gov.uk/government/news/higher-standards-as-thousands-more-take-gcses-at-the-right-time

DES. (1982). *Mathematics counts: Report of the committee of inquiry into the teaching of mathematics in schools under the chairmanship of Dr W. H. Cockcroft*. London: HMSO.

Dowling, P. (1998). *The sociology of mathematics education: Mathematical myths/ pedagogic texts*. London: Falmer.

Fairclough, N. (2001). *Language and power* (2nd ed.). Harlow: Longman.

Fisher-Hoch, H., Hughes, S., & Bramley, T. (1997). *What makes GCSE examination questions difficult? Outcomes of manipulating difficulty of GCSE questions*. Paper presented at the British Educational Research Association Annual Conference, York. Retrieved from http://www.cambridgeassessment.org.uk/ca/collection1/digitalAssets/113957_What_Makes_GCSE_Examination_Questions_Difficult.Outcomes_of_.pdf

Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal*, 28(4), 567–590.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Halliday, M. A. K. (1998). Things and relations: Regrammaticising experience as technical knowledge. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourse of science* (pp. 186–235). London: Routledge.

Halliday, M. A. K. (2003). *On language and linguistics*. London: Continuum.

Halliday, M. A. K., & Mathiessen, C. M. I. M. (1999). *Construing experience through meaning: A language–based approach to cognition*. London: Continuum.

Herbel-Eisenmann, B., & Wagner, D. (2007). A framework for uncovering the way a textbook may position the mathematics learner. *For the Learning of Mathematics*, 27(2), 8–14.

Hodge, R., & Kress, G. (1988). *Social semiotics*. Cambridge: Polity Press.

Hodgen, J., Marks, R., & Pepper, D. (2013). *Towards universal participation in post-16 mathematics: Lessons from high-performing countries*. London: Nuffield Foundation.

Kilpatrick, J. (2014). From clay tablet to computer tablet: The evolution of school mathematics textbooks. In K. Jones, C. Bokhove, G. Howson, & L. Fab (Eds.), *Proceedings of the International Conference on Mathematics Textbooks Research and Development (ICMT)* (pp. 3–12). Southampton: Southampton Education School, University of Southampton.

Kilpatrick, J. (2015). History of research in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 267–272). Berlin: Springer.

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.

Lemke, J. L. (1993). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing.

McSmith, A. (2007). Reading and maths standards falling in Britain says OECD. *The Independent*. Retrieved from http://www.independent.co.uk/news/education/education-news/reading-and-maths-standards-falling-in-britain-says-oecd-762805.html

Morgan, C. (1998). *Writing mathematically: The discourse of investigation*. London: Falmer.

Morgan, C. (2006). What does social semiotics have to offer mathematics education research? *Educational Studies in Mathematics*, 61(1/2), 219–245. doi:10.1007/s10649-006-5477-x

Morgan, C. (2009). Questioning the mathematics curriculum: A discursive approach. In L. Black, H. Mendick, & Y. Solomon (Eds.), *Mathematical relationships in education: Identities and participation* (pp. 97–106). London: Routledge.

Morgan, C. (2014a). Beyond communication: Using language for researching curriculum, pedagogy and policy in mathematics education. *Annals of the Polish Mathematical Society 5th Series: Didactica Mathematicae*, 36, 33–56. doi:10.14708/dm.v36i0.690

Morgan, C. (2014b). Understanding practices in mathematics education: Structure and text. *Educational Studies in Mathematics*, 87(2), 129–143. doi:10.1007/s10649-013-9482-6

Morgan, C., Tang, S., & Sfard, A. (2012). Doing the same mathematics? Exploring changes over time in students' participation in mathematical discourse through responses to GCSE questions. *Informal Proceedings of the British Society for Research into Learning Mathematics*, 32(3), 138–143. Retrieved from: http://www.bsrlm.org.uk/IPs/ip32-3/BSRLM-IP-32-3-24.pdf

Nachlieli, T., & Tabach, M. (Eds.). (2016). *Communicational perspectives on learning and teaching mathematics*. Special issue of *Educational Studies in Mathematics*, 91(3).

Newton, J. A. (2012). Investigating the mathematical equivalence of written and enacted middle school Standards-based curricula: Focus on rational numbers. *International Journal of Educational Research*, 51–52, 66–85.

O'Halloran, K. L. (2004). Discourses in secondary school mathematics classrooms according to social class and gender. In J. A. Foley (Ed.), *Language, education and discourse: Functional approaches* (pp. 191–225). London: Continuum.

O'Halloran, K. L. (2005). *Mathematical discourse: Language, symbolism and visual images*. London: Continuum.

Park, J. (2016). Communicational approach to study textbook discourse on the derivative. *Educational Studies in Mathematics*, 91(3), 395–421.

Paton, G. (2014). Drop in headline GCSE results 'reflects toughening up of standards', says minister. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/education/educationnews/11180601/Drop-in-headline-GCSE-results-reflects-toughening-up-of-standards-says-minister.html

Peled-Elhanan, N. (2015). *Palestine in Israeli school textbooks: Ideology and propaganda in education*. London: Tauris Academic Studies.

Pollitt, A., Hughes, S., Ahmed, A., Fisher-Hoch, H., & Bramley, T. (1998). *The effects of structure on the demands in GCSE and a level questions*. Cambridge: OCR.

Pons, X. (2012). Going beyond the 'PISA Shock' discourse: An analysis of the cognitive reception of PISA in six European countries, 2001–2008. *European Educational Research Journal*, 11(2), 206–226.

QCA. (2006). *Review of standards in mathematics: GCSE 1999–2004 and A level 1998–2004*. London: Author.

Rap, S., & Blonder, R. (2016). Let's face(book) it: Analyzing interactions in social network groups for chemistry learning. *Journal of Science Education and Technology*, 25(1), 62–76.

Reform. (2005). *Standards of public examinations in England and Wales*. Retrieved from http://www.reform.uk/wp-content/uploads/2014/10/standards_of_public_examinations_england_wales.pdf

SCAA. (1996). *Standards in public examinations 1975 to 1995*. London: Author.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22, 1–36.

Sfard, A. (2008). *Thinking as communicating: human development, the growth of discourses, and mathematizing*. Cambridge: Cambridge University Press.

Sfard, A. (Ed.). (2012). Developing mathematical discourse: Some insights from communicational research. Special issue of *The International Journal of Educational Research*, Vol. 51–52.

Sfard, A. (2013). Not just so stories: Practicing discursive research for the benefit of educational practice. In V. Farnsworth & Y. Solomon (Eds.), *What works in education? Bridging theory and practice in research* (pp. 139–150). London: Routledge.

Sfard, A., & Lavie, I. (2005). Why cannot children see as the same what grown-ups cannot see as different? - Early numerical thinking revisited. *Cognition and Instruction*, 23(2), 237–309.

Smith, C., & Morgan, C. (2016). Curricular orientations to real-world contexts in mathematics. *Journal of Curriculum Studies*, 27. doi:10.1080/09585176.2016.1139498