**Open-access writing: an investigation into the online drafting and revision of a research article in pure mathematics**

ESP research has provided an account of research articles (RAs) across disciplines using both text-analytical methods and ethnographically-oriented approaches. This study explores what additional insights are gained into the genre from the study of a collaboratively produced RA in pure mathematics, negotiated via an open-access research blog. The data consists of 659 thread comments posted by blog participants as they engage with the research and writing up process. Facets of research-based writing that preoccupy the blog participants are revealed, as well as how decisions pertaining to genre and dissemination outlets are made. In addition, blog posts point to how the RA is adjusted to cater for the more diverse readership that open-access knowledge dissemination may entail. The findings provide support for results of existing genre analyses of RAs in pure mathematics, and offer new insights into writing for publication practices in the discipline. Potential pedagogical applications of the findings are proposed.

**Key words**: blogs, disciplinary discourse, pure mathematics, writing for publication

## 1. Introduction

   Academia has now entered the blogosphere (Mauranen, 2013); although still in its infancy, the academic blog is used for knowledge dissemination, particularly within the hard sciences, enabling scholars to engage with academic communities, present work in progress, and receive feedback (Luzón, 2012). Thus, blogging can be seen as a new medium for performing tasks which academics already do; however, unlike traditional mediums of academic communication facilitated via established genres, academic blogs are not necessarily accessed by a clearly defined discourse community, and therefore constitute a potentially powerful means of reaching both scholarly and lay audiences simultaneously (Luzón, 2013a).

Research into the discourse, visual characteristics, operational facets and pedagogical applications of blog communication has started to accumulate (e.g. Gurak et al. 2004). Nonetheless, Myers' (2013) has pointed out the need for more qualitative studies into academic communication using digital media. Most ESP research on blogs to date has analysed the discourse features of posts and comment threads. For example, Luzón (e.g. 2011, 2012, 2013a, 2013b) investigates the interactional dimension of blog discourse, and how scientific knowledge can be repackaged to address a diversified blog audience. Theoretical questions have also been raised, such as whether generic status can be accorded to texts transmitted via this medium (e.g. Giltrow & Stein, 2009; Mauranen, 2013; Myers, 2010; 2013), the applicability of "discourse community" (Swales, 1990, p. 23) to describe a blogging community (e.g. Barton & Lee, 2013; Luzón, 2011; Mauranen, 2013; Schmidt, 2007), and what blogs reveal about genre change and recurrence (Miller & Shepherd, 2009).

Collaborative research blogs, although rarer than personal blogs maintained by individual researchers, offer additional scope for investigation; online research cooperation between scholars via blogs renders research and writing practices more visible, revealing how knowledge is constructed as well as reported as a final product. In other words, the affordances of digital media have the potential to make writing for publication practices more transparent (Myers, 2010). This offers an intriguing shift in perspective; some ethnographically-oriented investigations into writing for publication include Knorr-Cetina's (1981) investigation into the production of an RA, Lillis and Curry's (2010) research into the multilingual scholars' writing practices, Myers' (1990) account of the textual construction of scientific knowledge, and Bazerman's (1988) exploration of the development of the experimental report. Nonetheless, while research on L1 and L2 writing (e.g. Zamel, 1982) shows that writing academic genres is a complex and "recursive" process (Swales, 1990, p. 220), ESP genre research (e.g. Lin & Evans, 2012; Stoller & Robinson, 2013) has typically

focused on the end product (i.e. the published RA) providing limited insight into the journey of the article in its construction and the rhetorical decisions made by authors en route. This is no doubt in part due to the significant complication that "research activities occur in relative privacy" (Bazerman, 1988, p. 200).

This study addresses the issue of visibility by investigating the online co-authorship and revision of an RA in pure mathematics, facilitated via the *Polymath* 8(a) open-access research blog. *Polymath* is a collaborative research project which unites mathematicians in solving open problems and reporting results using a blog and wiki as the main channels of communication. The blog is particularly interesting in that experts elicit the views of non-specialist participants in an attempt to make the RA in production as accessible as possible to a wider academic audience. Through an investigation of the "internally-moderated changes" (Swales, 1990, p. 117) negotiated online, I explore what can be learnt about disciplinary discourse and writing for publication practices in pure mathematics from the study of this collaborative research blog.

The main contribution of the study lies in the method; my findings reveal how an analysis of discussions conducted via research blogs can provide insights into a discourse community's engagement with article construction and the RA genre. This particular blog not only provides support for descriptions of the pure mathematics RA obtained from studies using text-analytical and interview methods, but also reveals effects of the digital medium on writing practices, and how the genre of the RA could be adapted to a more heterogeneous audience.

The article is organized as follows: Section 2 summarizes our current knowledge of the discourse of mathematics. Next, some background to the *Polymath* blog is given, and analytical procedures are described. The results of the blog analysis are set out in Section 4, and illustrative extracts from the threads are discussed. In Section 5, I consider what can be

ascertained about the discourse and writing practices of scholars through the investigation of a collaborative research blog.

## 2. Previous research into the discourse of mathematics

It has been established that the RA differs along disciplinary lines (e.g. Hyland & Tse, 2004, Hyland, 2005; Hyland, 2006). Numerous studies drawing on ESP genre theory demonstrate specificity across a number of disciplines (e.g. Bruce, 2009; Lim, 2010; Lin & Evans, 2012; Yang & Allison, 2004), with epistemological and social factors (e.g. Becher & Trowler, 2001; Bernstein, 1999) providing a rationale for the variation observed.

In terms of mathematics, O'Halloran's research (e.g. 2000, 2005, 2010) takes a functional, multi-modal perspective, while education scholars have investigated discoursal identity construction (Burton & Morgan, 2000), and genres used in mathematics teaching (Artemeva & Fox, 2011). However, from an ESP perspective, mathematics is a somewhat neglected discipline; Swales et al. (1998) examined imperatives in research level mathematics, and Shaw (2006) described the integration of mathematical code in pure mathematics, engineering, and physics texts. Four recent papers based on textual analysis and interviews (McGrath & Kuteeva, 2012; Kuteeva & McGrath, 2015; Graves, Moghaddasi & Hashim, 2013; 2014) have provided a more detailed description of RAs. An investigation into the overall RA structure (Graves et al., 2013) found that the IMRD structure is not adopted, but rather an "Introduction-Results" organization. Kuteeva & McGrath (2015) draw attention to the two argument structures that run through RAs: the mathematical argument, which consists of definitions, theorems and proofs, and the meta-mathematical argument, which comprises inter alia motivations and examples. This argument structure entails an overall article shape which differs significantly from Hill, Soppelsa and West's (1982) hour-glass representation of empirical articles.

The organization of sections in pure mathematics RAs appears to be relatively non-standardized. While theoretical RAs contain at least an opening section and a proof, a prototypical structure for the remainder of the article is difficult to establish (Kuteeva & McGrath, 2015). Graves et al. (2013) identify "complementary introduction" sections, containing some prototypical introduction moves (Swales, 1990) while Kuteeva & McGrath (2015) use the term "contextual background section" to describe sections distributed throughout the article. In particular, the ending of RAs is the least standardized; some terminate at the conclusion of the proof (indicated by the QED square notation), while others pose questions for future research. Prototypical conclusion sections (e.g. Swales, 1990) are rare in pure mathematics (Graves et al., 2013; Kuteeva & McGrath, 2015). Thus, the rhetorical structure of the RA is fluid, enabling the author to adapt the RA to the content and reader.

A corpus study based on Hyland's (2005) stance and engagement framework found a low number of hedges and attitude markers compared with both hard and soft disciplines, but higher than expected shared knowledge and reader references (McGrath & Kuteeva, 2012). Interview data suggested an urge to uncover the aesthetic quality of mathematics in the drive to achieve an elegant and natural simplicity to the results (McGrath & Kuteeva, 2012).

**3. Method**

The present study is a descriptive, intrinsic case study in that the focus is on the "particularity" (Stake, 2008, p. 121) of a single research blog with a view to contributing to the ESP community's understanding of writing for publication and the RA genre. The primary data source is the *Polymath* blog. In addition, I refer to the first full draft of the article posted on *arXiv* (a widely used on-line repository for RA pre-prints in mathematics). Insights from previous studies on the discourse of mathematics are also drawn on extensively in the interpretation of my findings.

*3.1 An overview of the Polymath 8(a) blog*

The aim of *Polymath* is to conduct mathematics research online. *Polymath 8(a)* was selected as it was the most current in the series at the time of writing. The project has multiple dissemination outlets; participants collaborate via an open-access blog, and maintain a wiki containing figures, tables, lecture notes and slides. *Polymath* has been reported on by journalists in popular science articles (e.g. Klarreich, 2013), and has provided data for research into the dynamics of online collaboration (Barany, 2010).

The blog is innovative in its approach to including non-specialist mathematicians in knowledge construction, writing, and dissemination. It should be noted that 'non-specialist' and 'non-expert' here refer to individuals who do not work within the relevant sub-field of mathematics, students of mathematics, and scholars from other disciplines.

In total, 39 participants posted comments. The proportion of expert to non-expert participants is difficult to establish. Some participants self-identify as non-experts in preambles to their comments. For example: "I'm asking as a curios (sic) non-mathematician" (5.61). One relatively active participant uses the name of a famous biologist, and it appears that most of the professional mathematicians use their own names or initials. Other participants' levels of expertise can be inferred to a certain extent based on frequency and content of their blog comments. User names also provide clues (for example, 'somedude' is unlikely to be part of the scholarly community, but could be). 'Anonymous' may identify multiple authors, but is treated as one author in the count. As in other blog investigations (e.g. Luzón, 2013; Mauranen, 2013), consent was not solicited from the blog participants as all the material used in this study is published on the net.

As far as can be ascertained by a reading of the blog comments, drafts of sections authored by participants are posted onto a wiki. Revisions are made directly onto the wiki draft and then reported on the blog, or proposed and discussed via comments. Unfortunately, drafts of

the RA at different stages of construction are not available for analysis. At the time of writing, the article is under review at a journal.

My analysis of the blog threads begins when the coordinator announces via a post the start of writing the RA. This post is the first of five blog posts and threads. The previous multiple threads which show the process of proving the theorem are not included in the investigation, as my interest lies in the construction of the RA rather than the process of achieving a mathematical result. In his subsequent four blog posts at the start of each thread, the coordinator summarises the work completed over the course of the previous thread, and suggests areas to focus on. All threads are available here: http://terrytao.wordpress.com/2013/08/17/Polymath8-writing-the-paper/. The article resulting from the collaboration (Polymath, 2014a) is available here: http://arxiv.org/abs/1402.0811. Throughout this article, examples from the blog threads are numbered, whereas extracts from the RA are lettered. Underlining indicates my added emphasis.

*3.2 Approach to the analysis of the blog threads*

The blog data comprises 659 comments (totalling 57105 words including mathematical symbols) posted in response to five blog entries. I began by reading all the comments on the five blog threads in chronological order so as to gain insight into the types of revisions proposed. Next, I coded the first thread. Each comment was coded descriptively, such as "pointing out a spelling mistake" or "suggesting a change to the wording of the theorem" (see Table 1). The codes were collapsed (Dörnyei, 2007) by clustering similar topics (Creswell, 2009). This clustering process was guided by Kuteeva and McGrath's (2015) description of a pure mathematics RA, as the focus of the study is research-based writing. The codes provided a template (Crabtree & Miller, 1999) which was used to categorize the remaining threads. This resulted in six categories (see below). While a more fine-grained categorization was produced in the early stages of the coding process, the complexity of a more detailed account

was unhelpful in revealing the important themes in the data. For example, Myers (2010) in his analysis of the production of a wiki article collapses his data into five categories. Thus, the analysis is best described as inductive and deductive, based on a reading of the content of the blog comments.

An overview of the categories is provided in Table 1. They are 'mathematical argument', 'meta-mathematical argument', 'expositional structure', 'propositional development', 'formal', and 'operational'. In particular, the categories 'mathematical argument' and 'meta-mathematical argument' require unpacking. As described in Section 2, the arrangement of a pure mathematics RA can be described as a dual argument (Steenrod, 1981; Kuteeva & McGrath, 2015). The mathematical argument comprises the steps involved in proving a result through logical reasoning, the basis for mathematical knowledge construction. This argument is conveyed predominantly (but not exclusively) using mathematical code via lemmas (dependent, auxiliary results which support the main result), propositions (minor results), theorems (main results), and definitions. The extract below taken from the *Polymath* RA provides an example:

(a) On the other hand, the sum

$$\sum_{i=1}^{k_0} \theta(n + h_i) - \log 3x$$

can only be positive if $n + h_i$ is prime for *at least* two indices $i = 1, \dots, k_0$. We conclude that, for all sufficiently large $x$, there exists some integer $n \in [x, 2x]$ such that $n + h_i$ is prime for at least two $i = 1, \dots, k_0$.

In addition, RAs incorporate a meta-mathematical argument. The meta-mathematical argument fleshes out and contextualizes the result(s), and provides motivation and background for various stages of the proof and the study as a whole. For example, niche-carving (Swales, 1990) builds the meta-mathematical argument in that a motivation for the

study is provided, as well as justification for the approach used. Below is an example from the *Polymath* RA:

> (b) Besides improving the bounds for gaps between primes, this paper also gives an independent confirmation of the validity of Zhang's work, since we do not invoke any of his results without giving a full proof.

In this example, the authors underscore their contribution by previewing findings which both support a previous claim by way of a new proof, and expand an established result (the bounds for gaps between primes).

Remarks providing meta-commentary such as how a result could be extended, literature review (c), or justification for the approach (d) also construct the meta-mathematical argument, as do examples that support and elucidate the various stages of the proof. Examples follow:

> (c) **Remark 2.9**. There are many variants of definitions of distribution estimates for arithmetic functions in arithmetic progressions in the literature, going back (at least) to the individual estimates of Hooley and Selberg for the divisor function.

> (d) **Remark 7.4.** The reason that r is taken to be slightly less than $N$ is to ensure that a diagonal term is manageable when the time comes to apply the Cauchy-Schwarz inequality.

For elaboration on the mathematical and meta-mathematical argument, and more authentic examples from RAs, I refer the reader to Kuteeva and McGrath (2015).

The category 'expositional structure' pertains to discussions governing the flow of information in the paper. 'Propositional development' covers comments reporting ongoing, new mathematical knowledge construction which may or may not be reported in the RA. 'Operational' comprises workload organization, issues with technology, and discussions

pertaining to publication. 'Formal' refers to typos, spelling errors, individual word choice, and formatting. Examples of coded comments are provided in the supplementary materials.

Table 1

**Table 1:** Overview of the categorization of the blog comments. The numbers in brackets following comments indicate the thread number (1-5) and comment number.

The numbers in brackets in Table 1 and throughout the article refer to the thread number (1-5) and comment number. Thus, (2.15) shows that the comment appears under the second blog post in the sequence of five blog posts, and is the fifteenth comment in the thread. The classification of comments into categories is not mutually exclusive, as some comments deal with more than one type of proposed revision. For example, the following comment was allocated 'operational' and 'expositional structure' categories.

> (1) I don't plan to touch other sections of the paper yet (not that there is much content to any of them at this stage) [**Operational].** One organisational issue regards how to deal with the van der Corput estimates. There are two places where we use a van der Corput bound: firstly for the "Deligne-free" Type I estimates (and also the Type II estimates), where we have to bound sums of the form [**Expositional structure**]. (1. 13)

Boundaries between categories can be fuzzy (e.g. Godsen, 2003; Hyland, 2005; Kuteeva, 2013; Luzón, 2013b; Mungra & Webber, 2010). Nonetheless, an example of the rationale applied in problematic cases is given below:

> (2) It may be helpful to add (perhaps in the introduction) a diagram showing the connections and implications among the important theorems (and perhaps also an index for nonstandard notations and definitions.) (2.18)

(3) I added a figure (Figure 1) indicating the general logical flow of the

argument, although I had to oversimplify a little bit. (2.19)

These comments pertain to flow of information in the article (and are thus tagged 'expositional structure'). In addition, they deal with information to be placed in the introduction, loosely functioning as Move 3 Step 3 of the CARS model (Swales, 1990). Therefore, the category 'meta-mathematical argument' was also allocated.

Following Luzón (2013a), the number of posts in which the strategies occurred and not the overall number of occurrences were counted. Of the 659 comments, 68 were excluded. These were comments which made typo corrections to previous comments, prank comments, and messages of congratulations and thanks. In addition, two extended discussions planning a popular article describing the experience of working on the *Polymath* project (*Polymath*, 2014b) and a follow-up research project (*Polymath 8b*) were excluded.

The coding was revisited several times. In addition, a mathematician was recruited as a rater. The mathematician had been an informant in a previous study which explored the argument structure of pure mathematics RAs (Kuteeva & McGrath, 2015), and was familiar with some of the categories used in this study. The advantages of collaborating with disciplinary informants in ESP research have been demonstrated (e.g. Hyland, 2005). The collaboration was particularly instructive in this case given the specificity of some aspects of mathematical discourse, which may have proven challenging for an applied linguist unused to dealing with mathematical code.

In order to test reliability, 15 comments were read by the informant. Across the 15 comments, 21 occurrences of categories were identified. Agreement on 17/21 occurrences of the categories was achieved initially. Discrepancies appeared to arise from the rater's reading of the comments in isolation rather than within the context of previous and subsequent

comments on the blog. With the provision of more contextual information and further discussion, agreement was reached on the remaining items.

## 4. Findings

The findings are presented as follows. Table 2 shows the number of occurrences of each category. Following the presentation of the numerical data, examples of the categories are discussed.

*4.1 Aspects of academic writing discussed via the blog*

Table 2 shows the total number of references to a given category in the blog comments.

**Table 2**: Number of occurrences of each category in comments across the five threads. The counting unit is a single blog comment. Note one unit may contain more than one feature.

Overall, categories which reveal less about mathematics discourse and writing for publication predominate. Given that the most frequent category is 'formal', it seems that participants were most preoccupied with issues such as typos and formatting. The second highest occurrence is 'mathematical argument'. The meta-mathematical argument received considerably less attention. While alterations to the dual argumentation occur throughout, in this case, the mathematical argument seems to require more comment. The expositional structure appears to be relatively uncontroversial in that specific reference to the flow of information occurs in only 35 comments. Conversely, 75 comments contained discussion pertaining to propositional development, suggesting that new mathematical knowledge production is ongoing during the writing and internal moderation process. These trends are discussed further in Section 4.2.

*4.2 Exploration of the categories*

In this section, each category is discussed in more detail, illustrated by extracts from the blog threads.

*4.3.1 The mathematical argument*

In total, 158 of the comments made reference to the mathematical argument. Extract 4 is relatively representative. Here, the author of a section justifies the chosen arrangement for the mathematical argument.

> (4) What I elected to do instead is to insert into the paper an older Type I estimate (…) this gives a slightly simpler proof of Zhang's theorem (and also the older Type I estimate can be used to motivate the more complicated Type I estimates we have). (2.13)

The justification for the revision can be interpreted in two ways. The first is cognitive scaffolding, where the decision is made to present a simpler proof before a more complex one. Alternatively, the writer could be foregrounding the more appealing result, as simplicity in a proof is highly valued by the disciplinary community.

High premium is also placed on the aesthetic quality of mathematics (e.g. Artemeva & Fox, 2011) in that a well-constructed proof is considered to have an aesthetically pleasing dimension, which, along with simplicity, can serve as a guide to reliability (McGrath & Kuteeva, 2012). The following observation taken from interview data with a mathematician (McGrath & Kuteeva, 2012, p. 168) encapsulates this view:

"This whole aesthetic appreciation is very potent in pure maths, it's a huge tradition. Some of us share the view that deep down, there is an aesthetic and that it is a very dependable guide that really proofs or lines of reasoning will be very be very beautiful and have an aesthetic appeal, and feel like the right way to do something.

In the blog comments, participants also refer to this aesthetic appreciation. In the following example, a proposed argument structure is described as "nicer-looking". While this description could be attributed to a lack of precise vocabulary to discuss rhetorical issues, the comment nonetheless invokes the aesthetic dimension:

> (5) It seems to me that Section 4.2 can be done asymptotically (…) using a variant of Lemma 4.2 that has a coprimality condition inserted (…) <u>This might even give a slightly nicer-looking argument</u>. (4.27)

Another consideration in the construction of the mathematical argument appears to be the ratio of written text to mathematical code. For example, from (6), we learn that "chunks of text" can be unappealing:

> (6) …this would be a more compact description than just listing the numbers, and it would make <u>it more easy for the reader to verify admissibility (...)</u>. But it would still be a large <u>ugly chunk of text</u>, so that's the consideration to be made here. (2.25)

It would seem then that there is a balance to be struck between how much space is allocated to running text rather than mathematical code in sections which build the mathematical argument.

Many comments in this category highlighted notational issues. Notation "creates shared knowledge between writer and reader from which the results arise" (Graves et al., 2013, p. 425) and comments on the blog reveal the importance of using notation that aligns with community expectations and values. In the following example, shared knowledge is disrupted as blog participants try to follow the notation used by previous authors (what would be described as "standard notation" in an RA). Conflict occurs as the two authors use the same notation for different purposes:

(7) Right now <u>we are using Δ for two different things</u>: firstly, following the notation of Zhang, we are using $\Delta(\alpha; a(q))$ to denote the (signed) discrepancy of $\alpha$ in the residue class $a(q)$. Secondly, following the notation of <u>Motohashi and Pintz</u>, we are using Δ to denote a certain multiplicative function related to the GPY sieve… (3.105)

Discussions pertaining to notation were also concerned with the visual aesthetic of the mathematical code. For example, in the following extract, a participant reports changing the notation in order to improve the aesthetic appeal:

(8) I changed this just for typographical reasons (especially, <u>I found the multiple \sqrt{\tilde{\theta}} in the Bessel functions somewhat ugly</u>…) (4.63)

Interestingly, notational issues continued to be remarked upon as late as Thread 4. This could be a result of the collaboration, but it is in any case an example of how the real research and writing up process diverges from what is reported in an RA (e.g. Swales, 1990); in pure mathematics, notation is prototypically set prior to the presentation of the proof. The following extract from the *Polymath* RA illustrates (note that 'Theorem' indicates the start of the proof):

(e) <u>We denote</u> by $j_v = j_{v,1}$ the first positive zero of the Bessel function $J_v$ (which is known to have infinitely many positive simple zeros; basic properties of the Bessel function may be found in [1]). The result of [19] is:

**<u>Theorem 2.12</u>** (Optimized Goldston-Pintz-Yildirim). [19] Let $k_0 \geq 2$ be a fixed integer, and let $0 < \bar{\omega} < 1/4$ be a fixed quantity such that

$$1 + \bar{\omega} > \frac{j_{k_0-2}}{k_0(k_0 - 1)}$$

Then $EH[\frac{1}{2}+2\varpi]$ implies $DHL[k_0, 2]$.

*4.3.2 The meta-mathematical argument*

As observed in Section 4.1, fewer comments pertaining to the meta-mathematical argument, "the complementary *informal* or *introductory* material consisting of motivations, analogies, examples, and meta-mathematical explanations" (Steenrod, 1973, p. 1) were found. Nonetheless, clear (and rare) examples of adjustments to the meta-mathematical argument appeared in discussions surrounding the introduction. In the following examples, I interpret the participants' discussion as pertaining to how to "create a research space" (Swales, 1990, p.140). In (9), the blog participants discuss what ESP analysts would term Move 2 Step 1D (continuing a tradition), and in (10), Move 3 Step 2 (announcing principal findings).

> (9) I've made some minor edits to the <u>abstract and introduction</u>, <u>to emphasise</u> more the distribution theorem on primes in arithmetic progressions we have <u>that improves upon Zhang's theorem (and also does not necessarily rely on Deligne's work</u>). (4.112)

> (10) <u>I have also been thinking about the best way to present the paper</u> (…) Indeed an <u>essential contribution of Polymath8</u> is, in my mind, <u>the new result</u>s on the distribution of primes in arithmetic progressions to large moduli (I think that the <u>new variants </u>of the exponent of distribution (…), <u>are certainly going to be useful in other applications</u>. I also think that (building on Zhang's work) <u>we bring a potentially critical new insight to this question</u>…. (4.113)

In these blog discussions, the cumulative knowledge making practices of the discipline are highlighted; rather than indicating a research gap, the authors underscore their contribution by showing how a previous result has been built upon, and the applicability of their work to

future research. The RA extract below shows how the contribution is framed in the introduction.

> (f) Our arguments moreover give <u>a new proof of Zhang's distribution theorem</u> on primes in arithmetic progressions (which improves upon the Bombieri-Vinogradov theorem when restricted to smooth moduli and to a single congruence class), which again does not necessarily rely on Deligne's work (…)

Interestingly, the applicability of the results (11) is not visible in the introduction. Instead, reference to these applications appears in other sections of the RA. This supports previous research (Kuteeva & McGrath, 2015; Graves et al., 2013) which found that rhetorical work prototypically placed in introductions in other disciplines can be distributed throughout sections of mathematics RAs. However, comments pertaining to niche-carving are scarce, suggesting either that the initial author of the section wrote a convincing introduction from the outset, or that highlighting a gap in the knowledge is not high priority (see Kuteeva & McGrath, 2015, p. 10).

Comments pertaining to the addition or alteration of meta-mathematical argumentation, such as in footnotes (11) or remarks (12), appear throughout the blog threads. In the following extracts (from threads 2 and 5), blog participants propose adding justification for the approach taken at various stages in the proof. Extracts (g) and (h) show how the comments appear in the RA.

> (11) Right before Theorem 3.1: I added a footnote on the recent results on bounded gaps between prime triples, quadruples, … as <u>an extra motivation</u> for having efficient methods to find narrow admissible tuples. (5.75)

(g) For the purposes of establishing bounded gaps between primes, <u>an even shorter and simpler argument is now given in [62]</u>;

(12) Yes, I agree that the current setup (based on the sheaf interpretation of the function itself rather than the Fourier transform) <u>is more conceptually natural</u> and better suited for future improvements; <u>I've added a remark about the alternate route at Remark 8.19</u>… (2.124)

(h) **Remark 8.19.** In the remainder of this paper, we will only use the bounds (8.16) and (8.17) from Theorem 8.17. <u>These bounds can also be expressed in terms of the Fourier transform</u> (…)

These additions demonstrate how mathematicians momentarily deviate from the mathematical argument in the text in order to provide justification and contextualisation for the mathematical choices made (McGrath & Kuteeva, 2012). This includes commentary on alternative but less "natural" avenues of mathematical reasoning (12) that were not pursued.

Judging the appropriate stance is of critical importance in writing for publication (e.g. Myers, 1985, cited in Swales, 1990; Hyland, 2005) and yet only one comment was found to discuss the level of knowledge claim in conjunction with an external result (this forming part of the meta-mathematical argument). In the following extract, a blog participant queries the use of the booster "invariably" (Hyland, 2005).

(13) At the end of Section 3.2, there is mention of something being <u>"invariably" true</u>: is it an experimental fact or a theorem? (3.97)

Therefore, unlike Knorr-Cetina's (1981) study, where biologists' claims were considerably weakened over the moderation process, the writers of this article do not discuss hedging. This

is perhaps to be expected, given that knowledge verification in mathematics is proof based, which makes hedging mostly redundant (McGrath & Kuteeva, 2012).

References are discussed in 12 comments across the five threads and are used for attested functions such as crediting other people's work and positioning (e.g. Harwood, 2008). Participants consult on which references to use, as they search for the most "standard" or "usual":

> (14) Do you have a standard reference for the various properties of coupon collector / balls in bins that are used for the heuristic analysis of the algorithm? Those I know don't really discuss, e.g., how many empty bins remain after a given number of steps, but it would be good to add a reference. (3.91)

> (15) My usual reference for this sort of thing is "Randomized Algorithms" by Motwani and Rhagavan. (3.92)

> (16) Sometimes attribute the prime tuples conjecture to Hardy-Littlewoord, and sometimes to Dickson-Hardy-Littlewood. Maybe we should uniformize this? Which one is more standard? (4.133)

The blog comments highlight that a demonstration of knowledge of the "standard reference" is an important rhetorical tool when writing for publication, and that "community recognized points of reference" are perpetuated through academic discourse (McGrath & Kuteeva, 2012, p. 169).

### 4.3.3 Structure

I now turn to the comments which discussed the RA structure. Structure refers to how the flow of information is regulated in the article, and therefore, to extracts which deal with whole

sections of the article, as well as placement of elements of the mathematical argument (i.e. lemmas, theorems etc.) or meta-mathematical argument. For example, the following extract pertains to the mathematical argument, but clearly also considers placement of the content:

> (17) I've added a lemma 4.10 which encapsulates the use of the Mertens formula for the sums over primes which occur twice in this section and two times more in the next. The location might not be the best: it's analogue to Lemma 4.2, and maybe would be better around there? (4.30)

It has been found that articles in pure mathematics do not adhere to a rigid framework, allowing the writer to adapt the flow of information to cater to the needs of the result and audience (Kuteeva & McGrath, 2015). This fluidity is exemplified as the authors seem to try to take into consideration the interests of a more heterogeneous readership. In (18), the coordinator states that the proposed structure facilitates selective reading:

> (18) I have tried to structure the paper so that the deepest arguments – the ones which rely on Deligne's theorems – are placed at the end of the paper, so that a reader who wishes to read and understand a proof of bounded gaps that does not rely on Deligne's theorems can stop reading about halfway through the paper. (Coordinator blog post 1)

The placement of sections is also determined by what is most logical in terms of supporting the argumentation. This is shown in the following comment where the participant reflects on where to place material which chiefly builds on the meta-mathematical argument:

> (19) I didn't want this section to overwhelm the rest of the paper since it is not part of the main argument (but rather an explanation as to why a certain

component of that argument is not expected to be improvable). <u>I'm still a little uncertain exactly what to do with this section</u>... (2.69)

Nonetheless, from what can be ascertained, the structure of the paper as was first planned does not alter significantly over the course of the writing. Indeed, most comments pertaining to structure occur in Thread 2. One noteworthy alteration to structure is proposed by a non-expert mathematician from the field of biology. In (20), the biologist proposes a section containing two prototypical conclusion moves: summarizing the results (inferred from the reference to "synthesis") and recommendation (Swales, 1990):

(20) I'm wondering whether there is a case for concluding the paper with a short section <u>summarising the current understanding of what avenues are available – and, critically, of how promising they seem – for further progress on H</u>. It seems to me (though I concede that this may be because my own field, <u>biology, may simply have different conventions for the structure of papers</u>) that the paper as it stands not <u>only lacks such a synthesis</u> but also is overall rather patchy in terms of describing which avenues are and are not promising. (1.87)

While the recommendation move has been found in pure mathematics RAs, a synthesis move does not appear to be prototypical (Kuteeva & McGrath, 2015). In this case, although the mathematician responds positively, we can interpret from his answer (21) that a more natural organization is the distribution of remarks throughout. Indeed, such a concluding section is not visible in the draft of the RA.

(21) <u>Hmm, that sounds like a good idea</u>; I've added a stub of a section on possible improvements to the end of the paper. <u>Of course we would also be inserting remarks at various junctures of the paper</u> when some further

optimisation looks likely or unlikely, but it does make sense to summarise all

that in one concluding section. (1.88)

The unfamiliarity with remarks from non-disciplinary specialists is again revealed in the

following comment from the biologist:

> (22)…perhaps <u>Section 11 could be the repository for all the various remarks</u> <u>and</u>
> <u>discussions</u> (some of them extensive, some probably very brief, <u>some of them</u>
> <u>already present elsewhere in the paper</u>, some not) about why this or that aspect
> of the current argument is likely to be rather refractory to further improvement.
> (2.70)

Rather than inserting remarks and discussion throughout the text, the biologist suggests

morphing the article into a structure more common to the natural sciences; by placing this

content together in one section, a structure that is more reflective of his own disciplinary

discourse would be achieved. Again, no Section 11 is visible in the RA draft. Instead, the

article terminates with the end of the proof (i), a relatively standard way to conclude an article

in pure mathematics (Kuteeva & McGrath, 2015):

> (i) It is now elementary to check that these give the bounds of Proposition 10.4
>
> (note that $m'y' = my$).

It would seem that structural conventions are not readily subverted if the aim is publication in

a disciplinary journal.

### 4.3.4 Propositional development

Propositional development pertains to new mathematical results that are produced and

discussed in tandem with the production of the RA. An example follows:

> (23) In order to estimate $G_x(0,0)$ we see from its definition (and lemma 4.5) that

(2) $G_x(0,0) \leq x J_{k_0-2}\left(j_{k_0-2}\sqrt{x}\right)^2$

whenever $j_{k_0-2}\sqrt{x} < j_{k_0-3}$.

The last condition is equivalent to

(2') $x < \left(j_{k_0-3}/j_{k_0-2}\right)^2$.

It is easy to verify that the last condition is equivalent to

(2'') $x < 1 - 2/k_0 + O\left(k_0^{-5/3}\right)$. (2.89)

As can be seen, the discourse of this category is at times indistinguishable from the discourse of a published RA. The mathematical code is integrated in the running script, and other attested features (McGrath & Kuteeva, 2012) such as frequent use of *we* and shared knowledge references (Hyland, 2005) are present.

The majority of the comments pertaining to ongoing research appear early in the discussions (Thread 1), which is probably to be expected. Nonetheless, research continues throughout the threads suggesting that the mathematical result presented to the reader is not necessarily finalized before writing begins. Indeed, new ideas are sparked and problems discovered during the writing process itself. For example, (24) and (25) reveal that a problem in the proof is discovered as late as Thread 5, prompting a hasty return to notebooks:

> (24) I checked the exponents of q_0 and there does not seem to be merely a computational mistake, so one has to find another workaround to get rid of this problem. I will continue thinking about it…(5.46)

> (25) Here is a possible fix. First observe that if the quantity H defined before Remark 7.7 (or in (10.1), which is essentially the same quantity) is less than 1, then $\hat{\Sigma}(b_1, b_2)$ vanishes and we have nothing to prove. So we may assume wlog that $H \geq 1$. (5.50)

It appears that a definition is tweaked so that the subsequent stages of mathematical argument hold. However, in the 'research story' of mathematics RAs, definitions are fixed prior to the unravelling of the proof. As has been shown to be the case in other disciplines (e.g. Swales, 1990), messy research stories are reformulated as a neat and sequential event when reported in the RA.

*4.3.5 Formal*

Comments pertaining to typos and spelling errors were by far the most frequent. Also common were issues with formatting, and one or two instances of problematic word choice were found. In this example, a native English speaker highlights a perceived problem with a non-canonical use of the word "apparition" (in the RA, "appearance" is used).

> (26) Two lines after (4.18): though I like the poetic imagery of "apparition", perhaps "appearance" would be better here. (In English, the word "apparition" has connotations of being ghostly.) (3.124).

The interest with revisions relating to these issues lies in their frequency. The constant need for extensive proofreading and correction is perhaps due to how parts of the text were produced (an amalgamation of blog comments), but also points to the intricacy of mathematics discourse. The frequency could also be attributed to the participation of non-expert mathematicians, who lack sufficient rhetorical and/or disciplinary knowledge to engage with other revision categories (e.g. Faigley & Witte, 1981; Sommers, 1980). Nonetheless, comments on all aspects of the article from all participants were encouraged by the blog coordinator:

> (27)…any feedback on the draft paper is encouraged, even from (or specially from!) readers who have been following this project on a casual basis, as this

would be valuable in <u>making sure that the paper is written in as accessible as</u>

<u>(sic) fashion as possible.</u> (Coordinator blog post 1)

An interesting discussion in this category deals with the issue of English variety. The matter is resolved simply by the adoption of American English, based on it being the majority variety among the blog participants.

> (28) On a related note is the question of whether to use <u>American English or</u>
> <u>Commonwealth English for the text</u>; given that (as far as I am aware) <u>I am the</u>
> <u>only active participant initially from a Commonwealth country, I am happy to</u>
> <u>defer here to the American English standard</u>. (2.8)

### 4.3.6 Operational

The final category corresponds to work actioned or completed, technology issues, and discussions about publication. A total of 120 comments made reference to one or more of these issues, around a fifth of the total number. While some comments dealt with project coordination, most simply reported that work had been completed, often by stating "done" in response to a previous post prompting revisions. The majority of these comments are of limited interest; nonetheless, a small number refer to journal publication, and provide insight into the pure mathematics RA. Surprisingly, the genre is not fixed at the start of the writing process:

> (29) I hadn't realised that the <u>net length of the paper</u> is likely to <u>exceed the 100</u>
> <u>page mark </u>(…). This may complicate the question of where to submit this paper
> (or whether to <u>convert this paper into a monograph</u>). (1.94)
> (30) Another possibility is Memoirs of the AMS, which is specifically focused
> on long monographs, <u>but perhaps we now have the opposite problem that our</u>
> <u>paper is a little short for a monograph</u>…" (2.64)

Unfortunately, there is no further discussion as to what conversion to a monograph entails. The RA posted on *arXiv* is around 160 pages, but nonetheless seems to be submitted to a standard journal, suggesting that length of text is not necessarily a genre determinant in pure mathematics. It is also apparent that the length of RAs in pure mathematics can greatly exceed the prototypical 10 000 words suggested by Swales (1990). One possible factor in this case may be the additional explanation included to ensure readability for the non-specialist audience, or it could simply reflect the complexity of the argument:

> (31) The size should not be a major issue for the journal, so I would not seek to artificially shorten the paper at the expense of readability. (Coordinator blog post 3)

Comments in this category also revealed an innovative approach to dealing with the issue of authorship. The following comment proposes a policy of self-selection according to quality of contribution to the project:

> (32) …people who feel that they have made a significant mathematical contribution to the project (of the level commensurate with that of a co-author in a traditional mathematical research paper) can add their own name, contact information, and grant information to the first section of the wiki page. Those who feel that they have made an auxiliary contribution to the project (e.g. stylistic suggestions, locating references, etc. – commensurate with a mention in the Acknowledgments section of a mathematical research paper) – can add their name to the second section of the wiki page. (1.141)

Authorship is defined by input into the mathematical result, irrespective of involvement in writing. This may be due to the digital medium; stretches of mathematical discussion on previous blog threads appear to have been cut and pasted directly into the article.

To conclude this section, it appears that in arranging the article, authors draw on cognitive and affective considerations, and adapt the flow of information to the needs of readers with varying interest levels. Notation generates discussion throughout the writing process, and aesthetic values come into play. Little discussion occurs around how to frame the argument in the introduction, but references are used to fulfil standard rhetorical functions. Mathematical knowledge construction continues throughout the writing process, and significant effort is expended on proof-reading and checking, an activity that is ongoing throughout.

## 5. Discussion and conclusion

The aim of this study has been to explore whether the study of the *Polymath* collaborative research blog sheds light on behind-the-scenes activities pertaining to writing for publication in pure mathematics, and the RA genre. Admittedly, the insights gained are not on the scale of ethnographically-oriented studies (e.g. Knorr-Cetina, 1981; Lillis and Curry, 2010; Myers, 1990), or textography (Swales, 1998). This is in part due to the fact that the data set is confined to blog comments and the resulting RA, without access to the various iterations of the article in progress or its journey through the review process. Furthermore, the findings cannot be generalized (in a quantitative sense) to all research-based writing in pure mathematics. Indeed, the case is somewhat idiosyncratic, both in terms of the blog forum and the relatively heterogeneous character of the participating community. Nonetheless, the resulting RA was written principally by members of an academic discourse community, with the aim of reporting a mathematical result derived from proof in a scholarly journal.

As this study is, to my knowledge, unique in ESP research in that it explores what can be learnt about writing for publication through the analysis of blog discussions, rather than the discourse of blogs per se (e.g. Luzón, 2011, 2012), it is difficult to place the results within the context of other research. However, this investigation does provide support for previous

results arising from genre analysis of mathematics RAs using alternative data and methods (e.g. Graves et al., 2013; 2014; Kuteeva & McGrath, 2015; McGrath & Kuteeva, 2012). For example, while the majority of the comments on the blog pertain to checking for typos and collaboration coordination, other comments (17-19) provide evidence of the fluidity in the structure of the RA (Graves et al., 2013; Kuteeva & McGrath, 2015), a focus on the aesthetic, a balance between explanatory text and mathematical reasoning (5-7) (McGrath & Kuteeva, 2012), and the presence of a dual argumentation structure (Kuteeva & McGrath, 2015).

My study also explored the influence of digital media on the discourse of the discipline. Unlike external journal peer-review, feedback on the article from non-experts is explicitly elicited and facilitated by the open-access blog. With this more heterogeneous audience in mind, the authors adapt the text by pacing the flow of information (18), and by providing additional explanation (31). This seems to result in a (more) accessible (in the eyes of one mathematician), albeit lengthy research article:

> (33)… I also liked how each of the <u>arguments was worked out explicitly</u>. The <u>first few sections</u> were an <u>easy read for non-experts</u>. (4.121)

To conclude, it has been claimed that the research article has now been "thoroughly analysed (…) from *all angles*" (my emphasis) (Mauranen, 2013 p. 8). Nonetheless, I argue that the present study adds another 'piece of the puzzle' in terms of our understanding of research article construction in mathematics. The analysis of collaborative research blogs such as *Polymath* has the potential to provide a new opening in EAP/ESP methodology that increases our knowledge of academic genres. Discussions surrounding production are rendered visible, enabling a virtual "non-participant observational approach" (Anna Mauranen, p.c).

Swales (1990, p. 130) observes that "the value of expert informants increases when they are conceived of as sources of information and insight, but also as objects of ethnographic study themselves as they negotiate textual material within their own environments." With this in mind, future research could combine the analysis of collaborative research blogs with established ESP research methods such as corpus and genre analysis, and interviews with expert informants. This may provide us with a fuller picture of disciplinary discourse and genres, and their evolution in the digital age.

The extent to which collaborative blogs are used by scholars when writing for publication is as yet unclear. Nonetheless, in research communities where they are used, there is potential pedagogical value; students learning to write RAs could be encouraged to follow or even participate in online discussions, in order to gain first-hand experience of article construction by professionals in their discipline. A student who followed the *Polymath* project commented that students rarely "get the chance to peak (sic) behind closed doors and watch professional mathematicians 'in the wild' like this" (Polymath, 2014b). While this student's focus may have been the mathematical content, an EAP teacher or subject specialist could draw attention to matters of academic writing, enabling students to become "ethnographers of genre" (Johns, 1997; Molle & Prior, 2008, p. 563). Examples from the blog could be used in writing seminars to generate discussion about the values (e.g. comments 6-8) and practices (e.g. 14-16) of discourse communities, and explored in conjunction with the resulting RA to motivate choices made by authors in terms of structure and positioning (e.g. 9, 10). Given the idiosyncratic authorship and construction of the article, the blog could also form the basis of critical discussions pertaining to power, conformity and innovation in research-based writing (e.g. 20, 21).

The results of the study provide further evidence that the construction of a pure mathematics RA is "neither simple, nor short, nor particularly natural" (Swales, 1990, p. 121). It is hoped

that the findings, combined with other recent results of genre analyses of RAs in the discipline, will alert EAP specialists to the specific needs of mathematicians, and inform genre-based needs analysis, syllabus and course design.

Finally, the *Polymath* blog is fertile ground for the exploration of digital academic communication. Researchers in the future may wish to take a more discoursal perspective, exploring issues such as identity construction among the participants, stance-making, and collaborative knowledge construction.

**References**

Artemeva, N. & Fox, J. (2011). The writing's on the board: The global and local in teaching undergraduate mathematics through chalk talk. *Written Communication, 28*(4), 345-379.

Barany, M. J. (2010). Polymath 1 and the modalities of 'massively collaborative mathematics'. *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*, http://dl.acm.org/citation.cfm?doid=1832772.1832786. Accessed 2014-04-24.

Barton, D., & Lee, C. (2013). *Language online*. Routledge.

Bazerman, C. (1988). *Shaping written knowledge. The genre and activity of the experimental article in science*. The University of Wisconsin Press.

Becher, T., & Trowler, P. (2001). *Academic tribes and territories*. Buckingham: Open University Press.

Bernstein, B. (1999). Vertical and horizontal discourse: An essay. *British Journal of Sociology of Education, 20*(2), 157-173.

Bruce, I. (2009). Results sections in sociology and organic chemistry articles: A genre analysis. *English for Specific Purposes, 28*(2), 105-124.

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

Crabtree, B., & Miller, W. (1999). Using codes and code manuals: A template organising style of interpretation. In B. Crabtree & W. Miller (Eds.*), Doing qualitative research* (2nd ed., pp. 163–177). Thousand Oaks: SAGE Publications.

Creswell, J. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks: SAGE Publications.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.

Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication, 32*(4), 400-414.

Giltrow, J. & Stein, D. (2009). *Genres in the internet: Issues in the theory of genre*. Amsterdam: John Benjamins.

Godsen, H. (2003). Why not give us the full story?: Functions of referees comments in peer reviews of scientific research papers. *Journal of English for Academic Purposes*, *2,* 87-101.

Gurak, A., Antonijevic,S., Johnson, L.Ratliff, C & Reyman, J. (2004). Into the blogosphere: Rhetoric, community, and culture of weblogs. Retrieved from http://blog.lib.umn.edu/blogosphere. Accessed 2015-03-01.

Graves, H., Moghaddasi, S., & Hashim, A. (2013). Mathematics is the method: Exploring the macro-organizational structure of research articles in mathematics. *Discourse Studies*, *15*(4), 421-438.

Graves, H., Moghaddasi, S., & Hashim, A. (2014). "Let G = (V,E) be a graph": Turning the abstract into the tangible in introductions in mathematics research articles. *English for Specific Purposes, 36*, 1-11.

Harwood, N. (2008). An interview-based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics*, *41*(3), 497-518.

Hill, S., Soppelsa, B., & West, G. (1982). Teaching ESL students to read and write experimental-research papers. *TESOL Quarterly, 16*(3), 333-347.

Hyland, K. (2005). Stance and Engagement: A model of interaction in academic discourse. *Discourse Studies*, 7, 173-192.

Hyland, K. (2006). Disciplinary differences: Language variation in academic discourses. In K. Hyland & M. Bondi (Eds.), *Academic discourse across disciplines* (pp. 17-45). Frankfurt: Peter Lang.

Hyland, K. & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, *25*(2), 156-177.

Johns, A. (1997). *Text, role and context: Developing academic literacies*. New York: Cambridge University Press.

Klarreich, E. (2013). Together and alone, closing the prime gap. *Quanta magazine.* Retrieved from: https://www.simonsfoundation.org/quanta/20131119-together-and-alone-closing-the-prime-gap/. Accessed 2014-04-24.

Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford: Pergamon Press.

Kuteeva, M. (2013). Graduate learners' approaches to genre-analysis tasks: Variations across and within four disciplines. *English for Specific Purposes*, *32*(2), 84-96.

Kuteeva, M. & McGrath, L. (2015). The theoretical research article as a reflection of disciplinary practices: The case of pure mathematics. *Applied Linguistics*, *36*(2): 215-235.

Lillis, T.M., & Curry, M.J. (2010). *Academic writing in a global context: The politics and practices of publishing in English*. London: Routledge.

Lim, J. M. (2010). Commenting on research results in applied linguistics and education: A comparative genre-based investigation. *Journal of English for Academic Purposes, 9*(4), 280-294.

Lin, L., & Evans, S. (2012) Structural patterns in empirical research articles: A cross disciplinary study. *English for Specific Purposes*, *31*(3), 150-160.

Luzón, M. J. (2011). 'Interesting Post, But I Disagree': Social presence and antisocial behaviour in academic weblogs. *Applied linguistics*, *32*(5), 517-540.

Luzón, M. J. (2012). "Your argument is wrong": A contribution to the study of evaluation in academic weblogs. *Text and Talk, 32*(2), 145-165.

Luzón, M. J. (2013a). Public communication of science in blogs: Recontextualizing scientific discourse for a diversified audience. *Written Communication*, *30*(4), 428-457.

Luzón, M. J. (2013b). 'This is an erroneous argument': Conflict in academic blog discussions. *Discourse, Context & Media*, *2*(2), 111-119.

Mauranen, A. (2013). Hybridism, edutainment, and doubt: Science blogging finding its feet. *Nordic Journal of English Studies, 13*(1*),* 7-36.

McGrath, L. & Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes, 31* (3), 161-173.

Miller, C., & Shepherd, D. (2009). Questions for genre theory from the blogosphere. In Giltrow, J. & Stein, D. (2009). *Genres in the Internet: Issues in the theory of genre.* Amsterdam: John Benjamins.

Molle, D. & Prior, P. (2008). Multimodal genre systems in EAP writing pedagogy: Reflecting in a needs analysis. *TESOL Quaterly*, *42*(4), 541-566.

Mungra, P., & Webber, P. (2010). Peer review process in medical research publications: Language and content comments. *English for Specific Purposes*, *29*, 43-53.

Myers, G. (1985). Texts as knowledge claims: The social construction of two biology articles. *Social Studies of Science*, *15*, 593-630.

Myers, G. (1990). *Writing biology: Texts in the construction of scientific knowledge.*

University of Wisconsin Press.

Myers, G. (2010). *The discourse of blogs and wikis*. London, UK: Continuum.

Myers, G. (2013). *Working and playing on science.* Retrieved from:

twitter.http://www.lancaster.ac.uk/fass/events/twitter_and_microblogging/video/GregMyers.h

tm. Accessed 2014-04-24.

O'Halloran, K. (2000). Classroom discourse in mathematics: A multisemiotic

analysis. *Linguistics and Education*, *10*(3), 359-388.

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

London and New York: Continuum.

O'Halloran, K. (2010). *The Semantic Hyperspace: Accumulating mathematical*

*knowledge across semiotic resources and modalities*. In F. Christie & K. Maton. (Eds).

Disciplinarity: Functional linguistic and sociological perspectives (pp. 217-236). London:

Continuum.

Polymath. (2014a). New equidistribution estimates of Zhang type, and bounded gaps between

primes. Retrieved from: http://arxiv.org/abs/1402.0811. Accessed 2015-06-23.

Polymath, D.H.J. (2014b). The bounded gaps between primes polymath project - a

retrospective. Retrieved from: http://arxiv.org/abs/1409.8361. Accessed 2015-06-22.

Schmidt, J. (2007). Blogging practices: An analytical framework. *Journal of Computer-*

*Mediated Communication, 12*(4), 1409-1427.

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers.

*College Composition and Communication*, *31*(4), 378-388.

Shaw, P. (2006). Relations between text and mathematics across disciplines. In. K. Hyland

and M. Bondi (Eds.). *Academic discourse across disciplines* (pp.103-122). Peter Lang.

Stake, R. E. (2008). Qualitative case studies. In N.K. Denzin & Y. S. Lincoln (Eds.),

*Strategies of Qualitative Enquiry* (pp. 119-149). Thousand Oaks: Sage Publications.

Steenrod, N. (1973). Introduction. In N. Steenrod, P. Halmos, M. Schiffer, and J. Dieudonné

(Eds.), *How to write mathematics* (pp. 1-18). Providence: American Mathematical Society.

Stoller, F.L., & Robinson, M. (2013). Chemistry journal articles: An interdisciplinary

approach to move analysis with pedagogical aims. *English for Specific Purposes*, *32*(1), 45-

47.

Swales, J. M. (1990). *Genre analysis: English in Academic and research settings*. Cambridge:

Cambridge University Press.

Swales, J. M. (1998). *Other floors other voices.* NJ: Lawrence Erlbaum Associates.

Yang, R., & Allison, D. (2004). Research articles in applied linguistics: Studies from a

functional perspective. *English for Specific Purposes*, *23*(3), 264-279.

Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quaterly, 1*(2), 195-

209.