The Role of Preprints in Scholarly Communication: Complementing the Functions of Academic Publishing

Kraakman, Demi

The Role of Preprints in Scholarly Communication: Complementing the Functions of Academic Publishing

Demi Kraakman (s3689247)
Leiden University
Media Studies: Book and Digital Media Studies
MA Thesis
Supervisor: Dr. Fleur Praal
Word Count: 19.416

Table of Contents

Introduction

Preprints are versions of academic manuscripts that are published online, for example on a preprint server such as arXiv, and that have not been peer-reviewed (yet). Preprints have been steadily gaining popularity over the past few decades, yet the number of preprints that were posted saw a substantial increase during the COVID-19 pandemic. The reason for this was the speed with which knowledge could be disseminated through preprints since preprints do not require peer review. Peer review is the standard in academic publishing and takes up a lot of time which therefore slows down the spread of scientific information. Due to pressing demands for a COVID-19 vaccine, time was precious and researchers opted to make their findings public as preprints instead of formally publishing them, in order to give other researchers quicker access to their findings and research. Of course, without peer review, these preprints do lack a certain level of rigorous quality control, which does come with drawbacks and concerns, but the time pressure of the COVID-19 pandemic overruled these concerns and prompted researchers to post their research as preprints regardless. Likewise, academic publishers became more lenient towards this, allowing preprints to be posted online due to the pressing circumstances. Academic publishers typically did not allow researchers to do so because they want to retain the exclusive right to publish the manuscript in their journal, therefore if researchers wish to publish with them, they cannot also post it as a preprint. The COVID-19 pandemic changed this.

Yet, the increased popularity of preprints does not appear to have lessened since the recent end of the pandemic and it is not expected that it will rapidly diminish either. Instead, this popularity has even spread to fields that have arguably had little to do with COVID-19 research. It would appear that the pandemic was the boost preprints needed to become more widely accepted across many academic fields. Now, many contemporary critics argue that preprints are becoming ‘a common part of the scholarly publishing process’ and are referring to preprints as an ‘important tool’ in academic publishing, as a way to ‘complement’ the current system.[1] However, academic publishing is a broad concept which includes many different functions, so which of these functions do preprints complement and how do they pose as an important tool? This thesis aims to provide an answer to this question. In order to do so, the functions of academic publishing will need to be identified.

[1] M.B. Hoy, ‘Rise of Rxivs: How Preprint Servers are Changing the Publishing Process’, *Medical Reference Services Quarterly*, 39 (2020), pp. 84-89 (p. 88); I. Puebla, J. Polka & O.Y. Rieger, *Preprints: Their Evolving Role in Science Communication* (California: ATG LLC (Media), 2021), p. 64; R.A. Gross, ‘Preprints and the Implications for Subsequent Peer Review and Publishing’, *Neurology*, 91 (2018), pp. 855-856 (p. 855).

Many scholars have written about the functions of (academic) publishing, but none of their theories fit well enough to be able to analyse preprints in relation to academic publishing. Some of the theories are too outdated to sufficiently analyse publishing during the current digital age, which is necessary because preprints are inherently digital documents spread on digital platforms. Others are too broad and focussed on publishing in general, rather than focussing on academic publishing. The reason why this thesis makes a distinction between academic publishing and more general publishing, such as trade publishing, is because the differences between the two are significant enough that it would influence the outcome of this thesis' analysis of preprints if it were analysed using general functions of publishing. Academic publishing has a different target audience and a different profit model, among other things, than, for example, trade publishing. Publishing is an umbrella term encompassing many different varieties of publishing, including academic publishing, all with slight differences that set them apart from other kinds of publishing. Since preprints relate to academic publishing specifically, this thesis will focus on establishing a theoretical framework that fits this specific type of publishing as to obtain the most accurate results from the analysis in this thesis.

So, in order to establish the functions of academic publishing that fit within the digital age and are specifically relevant to academic publishing, this thesis will compare five separate theories in order to formulate a new framework with which preprints and their importance as a complementary tool can be determined.

The first theory this thesis will draw from is Bhaskar's *The Content Machine*, which analyses publishing in general. Bhaskar is very inclusive with his definition of publishing, including even CDs and videotapes as products of publishing. Consequently, this theory is too broad to carelessly adopt the functions of publishing he identifies and use them for the analysis of preprints. So, while this theory is informative and useful, which is why it is included in this thesis, other theories are needed to build a comprehensive framework to analyse preprints specifically.

Likewise, Clark and Phillips' *Inside Book Publishing*, is also a broader theory focussed on publishing in general. While this theory is less broad than Bhaskar's and while it does explicitly address academic publishing in multiple subsections, the functions of publishing it identifies pertain to publishing in general and not academic publishing specifically. As a result, this theory does not fit well enough with the objective of this thesis to use it for its analysis. However, it does provide an additional perspective which serves to make the eventual theoretical framework more nuanced and comprehensive.

Then Origgi and Ramello's article, 'Current Dynamics of Scholarly Publishing', published in 2015, provides an outsider's perspective on the functions of academic publishing. Ramello is a professor of Industrial Economics and Origgi is an Italian philosopher who focusses on social sciences and epistemology in relation to new technologies and social media. Origgi has written an entire book on the influence of reputation on several aspects of life, [2] which also comes forth in her shared theory on the dynamics in academic publishing. With their outsiders' perspectives, Origgi and Ramello approach academic publishing, which is why this theory is also included. Due to their unrelated background, their insights can be valuable in order to establish a more nuanced framework that includes multiple perspectives rather than just perspectives belonging to scholars who are heavily involved in the publishing industry.

Following this theory, this thesis examines Roosendaal and Geurts' 'Forces and Functions in Scientific Communication: An Analysis of Their Interplay', which was a paper presented at a conference in Germany in 1997. This paper addresses scientific communication, which academic publishing is a part of, meaning that it relates more specifically to the functions of academic publishing than the theories of Bhaskar and Clark and Phillips do. Roosendaal and Geurts not only identify functions of academic publishing but also address how they intersect and relate to one another using a visualization of two axes to illustrate the connections between these functions. This paper provides a helpful perspective on the functions of academic publishing, but due to it being presented it 1997, it does not account for the technical innovations of today. Consequently, it does not fit well enough with the current objective to analyse preprints, since these are a typical phenomenon of the current digital age.

Lastly, this thesis draws on the theory of Anderson in *Scholarly Communication: What Everyone Needs to Know*. In this work on scholarly communication, Anderson also dedicates one chapter to academic publishing and the functions of it. Overall, Anderson presents a well-rounded, contemporary theory on the functions of academic publishing, which could be used to analyse preprints, yet this theory has been simplified and is not as nuanced as some of the other theories that this thesis analyses. *Scholarly Communication* is accessible and understandable for everyone, even those not all that familiar with the field, which also means that some of the concepts Anderson discusses are described in a brief and simple manner. While this approach is helpful, it implies that some subtleties and deeper facets of these

[2] G. Origgi, *Reputation: What It Is and Why It Matters*, (Princeton: Princeton UP, 2017).

functions may be overlooked. Consequently, incorporating the various other theories mentioned alongside this one can enrich the thesis by providing a more intricate and comprehensive framework.

Prior to the comparative analysis of these five theories, the first chapter will focus on conceptualizing preprints. It will provide a succinct definition of preprints and delve into the nature of preprint servers. Additionally, an exploration of the historical background of preprints will ensue, further elucidating how the emergence of the COVID-19 pandemic propelled preprints to the level of popularity that it has reached and the state of preprints now that the pandemic has ended. Subsequently, this chapter will present an examination of the advantages associated with preprints, along with an exploration of the drawbacks and concerns associated with the usage of preprints. Moreover, this first chapter will acknowledge the disciplinary differences, as not every academic field has embraced the adoption of preprints thus far.

The second chapter will contain the comparative analysis of the five aforementioned theories. Each of these theories will be further elaborated on, facilitating a deeper understanding of their perspectives on what the functions of (academic) publishing are. Throughout this chapter, the theories will already be compared and contrasted critically, culminating in a comprehensive evaluation of their merits and limitations. In the last subsection, these evaluations will come together to formulate a theoretical framework identifying the functions of academic publishing with which preprints can be analysed since it is applicable to both the digital age and academic publishing specifically.

With this theoretical framework established and preprints conceptualized, the third chapter will focus on analysing preprints and the role they have in relation to academic publishing functions. The objective of this analysis is to provide an answer to the question which of the functions of academic publishing do preprints complement and how do they pose as an important tool? By also drawing on an example from practice, this thesis aims to determine the extent to which preprints can have the relevant, complementary role that many scholars have claimed preprints have.

Currently, with the rise in popularity of preprints, much secondary literature is written about this topic and, as mentioned before, many scholars have started considering preprints a relevant part of scholarly communication. However, no other scholar has established a theoretical framework identifying the functions of academic publishing in order to analyse preprints and what aspects they complement. Many scholars have implied which functions preprints have a complementary or important role in, but this thesis aims to make these

implications explicit and substantiate them, thereby hoping to make a valuable contribution to the growing amount of literature on the topic of preprints.

Before commencing with the first chapter, I would like to establish that, in this thesis, the term 'academic publishing' will consistently be used to refer to the publication and dissemination of research and scholarly work. While the terms of 'scholarly publishing' and 'academic publishing' often overlap and are used interchangeably to refer to this process, this thesis has opted to use academic publishing for clarity's sake. Turmudi addresses this terminological issue in his article by illustrating how, depending on the source, different phrases and terms were used to refer to the same process.[3] He also specifically mentions how 'academic publishing' is used synonymously to 'scholarly publishing'.[4] Depending on the discipline, institutional practice or regional convention, there might be a preference for one term over the other, but it also happens that articles use the terms interchangeably, as exemplified by Texeira da Silva et al.'s article.[5] So, while acknowledging that there are some nuances to these terms depending on the context and preference, this thesis will use academic publishing to refer to the process of publishing and disseminating research and scholarly work.

[3] D. Turmudi, 'English scholarly publishing activities in the industrial revolution 4.0: What, why, and how?', *English Language Teaching Educational Journal,* 3 (2020), pp. 52-63 (p. 54).
[4] Ibid., p. 54.
[5] J.A. Texeira da Silva, et al., 'Predatory and Exploitative Behaviour in Academic Publishing: An Assessment', *The Journal of Academic Librarianship*, 45 (2019), pp. 1-8 (p. 6).

Chapter 1. Preprints

This thesis centres on preprints and their role in relation to academic publishing. In order to equip the reader with the necessary knowledge to fully engage the rest of this thesis, this chapter focusses on conceptualizing preprints. It commences by defining preprints and, in turn, will elaborate on preprint servers, which are the platforms on which preprints are posted.

Subsequently, this chapter will discuss the historical background of preprints, which will provide the reader with an understanding of the initial motivation behind the creation of preprints. Comprehending why preprints were created in the first place will give the reader a better idea of what motivates researchers to use preprints and the purpose it serves. Following the origin of preprints, the further development of preprints will be addressed, specifically the sudden boom in popularity of preprints due to the COVID-19 pandemic. To conclude this section, this thesis will elaborate on the current status and popularity of preprints.

Furthermore, this chapter will discuss both the advantages and the disadvantages of preprints in order to clarify why some researchers oppose preprints while others are proponents of using preprints. By addressing both sides of the debate surrounding preprints, this thesis will provide a more nuanced overview of the concept of preprints by not only focussing on its rising popularity but also emphasizing the drawbacks.

Lastly, the popularity of preprints differs significantly per discipline, which is why this chapter devotes an entire subsection on discussing this disparity. The disciplinary differences are important to acknowledge because it showcases that preprints have not been adopted in a uniform matter and may not be as beneficial to all disciplines.

1.1 Defining Preprints

Preprints are complete scholarly manuscripts that have not (yet) been through a formal peer review process and are made publicly available, often by uploading them to a preprint server, which can be accessed without any limitations.[6] Where some scholars upload their manuscript to a preprint server only, others upload a preprint while simultaneously submitting the

[6] Puebla, Polka, & Rieger, *Preprints: Their Evolving Role in Science Communication*, p. 1; 'Preprints in Europe PMC', *Europe PMC*, < https://europepmc.org/Preprints> (10 June 2023); and N.C. Penfold, & J.K. Polka, 'Technical and Social Issues Influencing the Adoption of Preprints in the Life Sciences', *PLoS Genetics*, 16 (2020), pp. 1-16 (p. 2).

manuscript to a journal for formal publication, meaning that the work is undergoing peer review while it is also already available as a preprint.[7] However, whether or not this is allowed depends on the publisher and their regulations. Publishers only accept manuscripts which have not been published elsewhere, meaning that, depending on whether or not they consider preprints a published work, they may not condone uploading a preprint.[8] Since preprints have become more popular, many academic publishers have explicit rules on whether uploading a preprint is allowed and there is a trend of publishers becoming more accepting towards preprints since the COVID-19 pandemic. This phenomenon will be discussed in depth further on in this chapter.

There are numerous preprint servers, some of which focus on one specific discipline, such as bioRxiv and chemRxiv, while others are interdisciplinary and label the manuscripts accordingly, such as arXiv. While some preprint servers only allow full manuscripts to be submitted, there are others who let scholars post preliminary results.[9] As stated above, there is no formal peer review for these manuscripts, but most preprints servers do provide basic screening, both to ensure that the manuscript is in accordance with the (disciplinary) scope of the server, labelled correctly, and to ensure it is a scientific manuscript which does not contain plagiarism.[10] The first preprint server was arXiv (pronounced 'archive'), which launched in 1991 and remains one of the biggest servers to this day.[11] Since then, many more preprint servers were created, such as bioRxiv, chemRxiv, and PeerJ Preprints, to name a few.[12] The functionalities of preprint servers can vary, but most allow revised version of manuscripts to be uploaded.[13] For example, when a manuscript has been accepted by a journal – if the academic publisher also allows this – this version may be uploaded and most servers provide the option to add a link to the final version, which is often called the 'version of record'.

By enabling free public access to preprints, these servers make a significant contribution to green Open Access, aligning seamlessly with the ethos of Open Access, which

[7] J.A. Teixeira da Silva, 'The Preprint Debate: What Are the Issues?', *Medical Journal Armed Forces India,* 74 (2018), pp. 162-164 (p. 162).
[8] For further reading on this topic see: T. Klebel, et al., 'Peer Review and Preprint Policies Are Unclear at Most Major Journals', *PLoS ONE,* 15 (2020), pp. 1-19.
[9] Puebla, Polka, & Rieger, *Preprints,* p. 49.
[10] 'Preprints', *Open Access Network,* < https://open-access.network/en/information/publishing/preprints#> (11 June 23).
[11] Teixeira da Silva, 'The Preprint Debate', p. 162.
[12] Unlike bioRxiv and chemRxiv, which are both disciplinary preprint servers, PeerJ Preprints is interdisciplinary and ranges from biology to computer sciences.
[13] Preprints', *Open Access Network,* (11 June 23).

advocates the free dissemination of scholarly knowledge.[14] Green Open Access is often referred to as self-archiving and means that a version of an author's manuscript that is also published by a publisher is posted in a repository or online.[15] Preprint servers are essentially repositories for scholarly manuscripts, which make them freely accessible, therefore they can contribute to green Open Access. However, not all preprints are also published by a publisher, so preprints only contribute to green Open Access to a certain extent.

Before a scholar decides to post their manuscript to a preprint server, most servers will offer them a choice regarding the licence they would want for their work. arXiv, for example, offers six Creative Commons licences to scholars who upload a manuscript, which give arXiv permission to post and distribute the scholarly work.[16] Aside from the CC-0 licence, under which a scholar gives up their copyright and puts their work in the public domain, the other five licenses allow the scholar to maintain the copyright of their intellectual work.

Once posted, a preprint is openly available to everyone and some preprint servers even offer a comment function for readers to interact with the preprint so authors can receive feedback from a larger audience.[17] This type of feedback can be especially useful with regard to preliminary results posted as a preprint, since researchers can improve their manuscript while still in the earlier stages of writing it.[18] However, not every preprint server offers a feedback function and it is no given that this function is indeed used. Moreover, because preprints are accessible to the public, there is little assurance that the person providing feedback is educated on the topic. Unlike formal peer review, the feedback on a preprint server does not necessarily come from an expert in the field, nor is it likely to be as thorough and critical as that provided through peer review. This is not to say that the feedback option on preprint servers cannot be beneficial, but it should be acknowledged that it is not comparable to peer review.

[14] Preprints', *Open Access Network,* (11 June 23); and Puebla, Polka, & Rieger, *Preprints: Their Evolving Role in Science Communication,* p. 55.
[15] Note that this does not have to be a definitive or edited version of the manuscript. See, for instance: 'What is Green Open Access?', *Author Services: Supporting Taylor & Francis Authors,* < https://authorservices.taylorandfrancis.com/choose-open/publishing-open-access/oa-green-gold/> (10 June 2023).
[16] 'arXiv License Information', *arXiv,* < https://info.arxiv.org/help/license/index.html>, (11 June 2023).
[17] 'Preprints', *Open Access Network,* (11 June 23); and Hoy, 'Rise of the Rxivs', p. 85.
[18] Puebla, Polka, & Rieger, *Preprints,* p. 49.

1.2 Historical Context: The Development of Preprints

Preprints, as we know them today, have a rich history that dates back to the 1960s. Although the first preprint server was launched in 1991, as mentioned before, a version of preprints had already been around for decades, predominantly in the fields of biology and physics.[19] The primary objective of sharing these forerunners of preprints was to share early findings with peers in order to receive feedback and gain new insights, thereby promoting collaboration. During the 1960s, scholars would circulate their preliminary work through the postal mail, sending it to other experts in the field. This practice originates from the field of biology and was soon thereafter adopted in the field of physics.[20] Physicists recognized the value of disseminating early-stage work, and they too began embracing these forerunners of preprints as a means of receiving feedback. As time progressed and technology advanced, the method of distribution also evolved. Due to the emergence of email, researchers began to adopt this as a method of exchanging their preliminary works.[21] Researchers could now send their preprints electronically, reducing the dependence on physical mail and enabling faster dissemination of their work. Then, in 1991, physicist Paul Ginsparg revolutionized the world of preprints by founding arXiv.[22] The online repository allowed scholars to reach a much wider audience with their initial findings. The advent of arXiv marked a turning point in the history of preprints as it now began to expand to more disciplines, including physics, economics, engineering, mathematics, finance, computer science, and statistics.[23]

However, notably, this expansion did not extend to biology and the life sciences, despite biologists pioneering the circulation of preliminary work among peers for feedback. Only after the launch of preprint platforms focused on biology and life sciences, such as bioRxiv and PeerJ Preprints in 2013, did uploaded preprints gain popularity within these fields.[24] The rising popularity of Open Science aided in the exponential growth in preprint's popularity, along with the initiatives that aimed to promote them.[25] In 2016, numerous more preprint servers were launched, which provided another boost in the popularity of preprints, although the level of acceptance still differed significantly per discipline, which was linked to

[19] Preprints', *Open Access Network,* (11 June 23).
[20] Ibid.
[21] Puebla, Polka, & Rieger, *Preprints,* p. 3.
[22] Ibid., p. 4.
[23] Hoy, 'Rise of the Rxivs', p. 85.
[24] Preprints', *Open Access Network,* (11 June 23); and Puebla, Polka, & Rieger, *Preprints: Their Evolving Role in Science Communication,* p. 5.
[25] Preprints', *Open Access Network,* (11 June 23); Puebla, Polka, & Rieger, *Preprints,* p. 6.

a culture of open science practices.[26] The more inclined a discipline is to share data, for example, the greater the adoption of preprints has been.[27] As more disciplines began using preprint platforms, they have slowly become more common and, as a result, preprints have been a growing practice ever since.

It was not until the global COVID-19 pandemic that preprints experienced a significant boost in popularity. This popularity stems from the need amidst the pandemic for rapid dissemination of knowledge and findings in order to find a vaccine, among other COVID-19 related research.[28] Formal peer review takes time in order to fulfil thoroughly and critically, which posed a problem during the global pandemic when time was of the essence. Therefore, many researchers began posting their findings as preprints in order to disseminate their results more rapidly. Additionally, many researchers were more inclined to share their work during the more preliminary stages with the objective of making their data available.[29] However, while this rapid dissemination of knowledge and data may have benefited collaboration amongst researchers and is likely to have made the vaccine possible as quick as it appeared, the sudden boom of preprints after the start of the COVID-19 pandemic has also raised many concerns. Scholars fear that the works published during this time may have been of lower quality due to the lack of peer review.[30] Moreover, academic journals shortened their peer review process of COVID-19 related manuscripts in order to publish them quicker, which might also have made the peer review processes less thorough and critical.[31] These concerns were later validated when many articles and preprints were retracted due to inaccuracies or misinformation, among other reasons.[32] This is especially concerning when taking into account that many policies were based on the information presented through unreviewed preprints and published articles with a shortened peer review time, since these could have later been retracted, meaning that these policies may have been based on false information.

Still, despite the concerns and drawbacks that accompanied this sudden boom of preprints during the pandemic, it cannot be denied that it has boosted the popularity of

[26] Puebla, Polka, & Rieger, *Preprints*, p. 6 and p. 8.
[27] Ibid., p. 8.
[28] Brierley, L., et al., 'Tracking Changes Between Preprint Posting and Journal Publication During a Pandemic', *PLoS Biology*, 20 (2022), pp. 1-22 (p. 1).
[29] Puebla, Polka, & Rieger, *Preprints*, p. 9.
[30] I. Kodvanj, et al., 'Publishing of COVID-19 Preprints in Peer-Reviewed Journals, Preprinting Trends, Public Discussion and Quality Issues', *Scientometrics*, 127 (2022), pp. 1339-1352 (p. 1340).
[31] Ibid., p. 1339.
[32] Y.V. Servryugina & A.J. Dicks, 'Publication Practices During the COVID-19 Pandemic: Expedited Publishing or Simply an Early Bird Effect?', *Learned Publishing*, 35 (2021), pp. 563-573 (p. 564).

preprints in general. While this increase in popularity was first only related to COVID-19 related research and the associated fields, the exponential increase in preprints has slowly spread to other disciplines.[33] Moreover, despite the end of the pandemic and there no longer being a pressing need for rapid dissemination of knowledge, the popularity of preprints has not decreased. In fact, its popularity appears to remain steady.

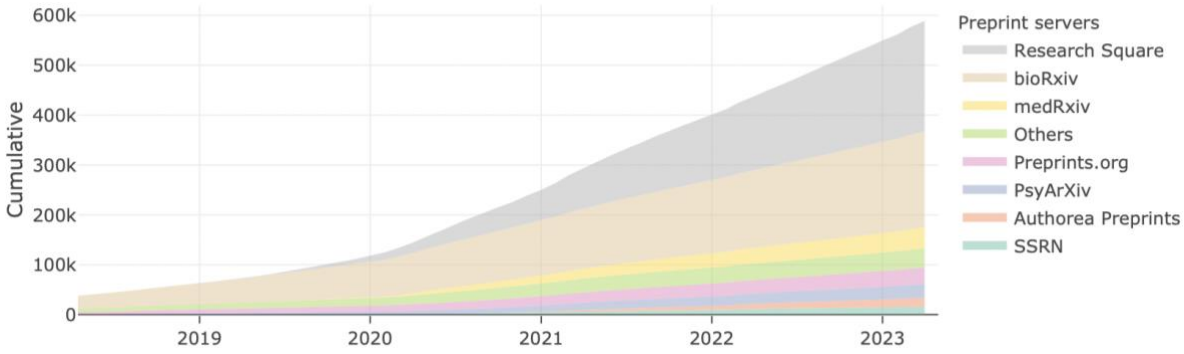

Figure 1: The Cumulative Total of Preprints per Preprint Server from April 2018 to April 2023.[34]

Figure 1 is a graph which shows the cumulative of preprints on each of the indicated preprint servers. The preprint server of Research Square is primarily focussed on life and social sciences. bioRxiv and medRxiv are variants of arXiv and focus on biology and health sciences respectively. Then the servers Preprints.org, Authorea Preprints, and SSRN are all interdisciplinary, and PsyArXiv is a preprint server specifically focussed on psychology. As can be seen in this figure, in early 2020, during the start of the global pandemic, there was a notable increase in preprint submissions to these preprint servers. Research Square and bioRxiv, both of which have a primary focus on life sciences, have experienced the most significant growth, which can be linked back to the pandemic and its effect on the adoption of preprints in this discipline especially. Furthermore, the exponential growth shown in this graph indicates that, even since the pandemic ended, there has not been a decrease in the preprints submitted to these servers, as the cumulative amount continues to rise steadily.

As a result of this rising popularity and use of preprints, academic publishers have gotten more involved with the phenomenon. Some publishers have acquired existing preprint servers, launched their own server, or partnered up with an existing server.[35] As mentioned in

[33] The reason for this will be discussed in subsection 1.5.
[34] 'Preprints in Europe PMC', *Europe PMC*, < https://europepmc.org/Preprints> (11 June 2023).
[35] Puebla, Polka, & Rieger, *Preprints*, p. 41-42.

the introduction, preprints are becoming a valuable part of scholarly communication and publishers are recognizing this and adapting to it in various degrees. Consequently, with academic publishers integrating preprints with the services they provide, it is logical that they allow, or even encourage, posting a preprint.[36] It is likely that they will encourage scholars to do so on their affiliated preprint server. For example, Elsevier bought SSRN in 2016 and they encourage their users to post a preprint on SSRN via their website.[37]

1.3 The Advantages of Preprints: What Makes Them Appealing to Scholars

While some of the advantages of posting preprints have already been touched upon briefly, it would be beneficial to discuss all of them in depth. This would provide a better understanding of what motivates scholars to make use of preprint servers. For this section, the main focus will be on the advantages of preprints for scholarly authors since the inception of preprints was made with authors in mind. The benefits of preprints for readers will also be touched upon, but most will revolve around authors. The way in which preprints can be advantageous for academic publishers will not be addressed here since this will be included in the third chapter when the complementary role and potential of preprints in regard to academic publishing will be discussed.

The most obvious advantage of posting scholarly work as a preprint is that of rapid dissemination, which was the main reason researchers posted COVID-19 related research as preprints during the global pandemic. Where publishing with a peer review journal may take months, posting a preprint makes the scholarly work immediately available.[38] Even without the time pressure of a pandemic, this rapid dissemination of scholarly knowledge is beneficial since it accelerates research, meaning that it increases the speed of new scientific discoveries.[39] Additionally, this is also an advantage for readers since it allows them to have quicker access to new research and developments. As mentioned before, this is one of the main advantages of preprints and a primary reason why scholars make use of them.

[36] W. Kaltenbrunner, et al., 'Innovating Peer Review, Reconfiguring Scholarly Communication: An Analytical Overview of Ongoing Peer Review Innovation Activities', *Journal of Documentation*, 78 (2022), pp. 429-449 (p. 436).
[37] 'SSRN', *Elsevier*, < https://www.elsevier.com/solutions/ssrn> (28 July 2023).
[38] Hoy, 'Rise of the Rxivs', p. 84.
[39] C. Strasser, 'Preprints: The Bigger Picture', *The Winnower*, 2016, pp. 1-6 (p. 1).

Another advantage that makes preprints attractive to both authors and readers is their accessibility. By posting on a preprint server, a scholar's work is exposed to a much wider audience than if it were published in a non-Open Access peer-reviewed journal, which is only read by those who can access the journal through licenses or subscriptions.[40] Preprints, on the other hand, can be accessed by anyone because they are made freely available to the public. Additionally, this promotes the transparency and reproducibility of research, especially when researchers share their earlier findings and acquired data as well. It is this kind of accessibility with no financial constraints that makes preprints appealing to both authors and readers.

This accessibility ties in with the increased visibility of a scholarly work. By making a work more accessible to a larger audience, it also increases the visibility of that work. This visibility is due to the fact that preprints are not posted on just any blog or website, but specifically on a preprint server, which contains multiple other academic works and research in the same discipline. Readers who use preprints as a resource to keep up with the latest developments in their field, which is a standard practice in disciplines such as physics,[41] will more easily come across a scholarly work if it is also posted as a preprint. This is especially the case for papers which have both been published in an academic journal and posted as a preprint. There are multiple versions of the publication findable online and considering how many preprint servers can provide preprints with a link to the version of record, this could indeed increase its visibility. Penfold and Polka also note this phenomenon that papers which have been preprinted garner more attention than those that have not, although they do point out that they have no concrete evidence to back this claim up.[42] Hoy, on the other hand, does provide statistics proving that this is indeed the case and that a paper which also has a preprint posted obtains more attention as well as more citations than articles published without a concurrent preprint.[43] Therefore, it appears to be a noticeable trend that a paper also posted as a preprint gathers more attention and visibility, which is therefore one of the advantages of posting a preprint.

Furthermore, while the lack of peer review is often considered a disadvantage of preprints – one which will be elaborated upon in the subsequent section focusing on the disadvantages – preprints do receive feedback which can benefit the scholarly work. First off, a scholar can receive feedback during the earlier stages of a research, prior to submission to

[40] Hoy, 'Rise of the Rxivs', p. 85.
[41] Penfold & Polka, 'Technical and Social Issues', p. 7.
[42] Penfold & Polka, 'Technical and Social Issues', p. 4.
[43] Hoy, 'Rise of the Rxivs', 86.

an academic publisher, which could benefit the overall quality of the work.[44] This was also the original objective of sharing preprints when they were still send to other scholars via postal mail. Of course, this feedback on a preprint server will not always be from an expert in the field, or even another scholar, but as Hoy points out, 'the aggregate knowledge of a large group can be a powerful tool'.[45] This type of feedback could still provide useful insights to a researcher and help advance the scholarly work. However, this would only work if readers actually use the feedback function, and whilst it appears that very few do, Penhold and Polka determined that a majority of the feedback given on preprints is provided outside of the preprint server.[46] A survey posted on bioRxiv established that over 40 percent of scholars who posted a preprint on this server received feedback on their preprint through other channels, such as social media, e-mail, or in-person discussions.[47] Therefore, while it may not always be through the means of the feedback function on a preprint server, preprints do invite feedback, which can help advance the overall quality of a scholar's research.

Another advantage of posting a preprint is that it can establish priority. Priority means that an author can claim ownership over intellectual property, therefore, when a scholar does a certain finding which has not been made before, this scholar holds the right to this idea. Preprint servers put a time stamp on each separate version of a preprint, which would ensure that an author can claim priority.[48] Especially considering the long peer review process required to get a paper published, a scholar might prefer posting it as a preprint in order to have a record to demonstrate that they were the first to establish a certain finding.

Finally, preprints could provide a citable record of research progress depending on the discipline, which scholars can then provide to funding agencies and promotion or hiring committees as evidence of their work.[49] In the current 'publish-or-perish' culture, in which a scholar's value is determined by their output in publications, preprint can offer a solution because they provide a tangible record of the scholar's work. These agencies and committees can trace a scholar's progress in real time through the use of preprints, so even when a scholar's work has not yet been published, they can demonstrate the advances that have been made.[50]

[44] Strasser, 'Preprints: The Bigger Picture', p. 1.
[45] Hoy, 'Rise of the Rxivs', 85.
[46] Penfold & Polka, 'Technical and Social Issues', p. 4.
[47] R. Sever, et al., 'bioRxiv: The Preprint Server for Biology', *bioRxiv*, 2019, pp. 1-19 (p. 8).
[48] Puebla, Polka, & Rieger, *Preprints*, p. 12.
[49] Ibid., p. 12.
[50] Teixeira da Silva, 'The Preprint Debate', p. 163.

This subsection has highlighted some of the key advantages of preprints, although it is likely that there are additional benefits that have not been fully explored here. By addressing these advantages, the popularity and value of preprints is more easily understood.

1.4 The Disadvantages of Preprints: The Concerns and Challenges

In order to provide a nuanced overview of the concept of preprints, this thesis will not only highlight the advantages, but also stress the disadvantages of preprints. By doing so, this thesis aims to give a more well-rounded perspective on the the multifaceted nature of preprints. Some of the disadvantages that will be addressed in this subsection have already been mentioned before but will be further elaborated on here. Like the subsection about the advantages, this subsection will also primarily focus on the drawbacks for academic authors, while occasionally also including the disadvantages for readers.

One significant drawback, as mentioned previously, is the question regarding the effectiveness of feedback received by preprints. Although preprints encourage and invite feedback, which undoubtedly proves to be valuable on occasion, evaluating the level of expertise of these commenters remains a challenge.[51] The additional effort required to screen every individual who provides feedback or comments often becomes a strenuous task. Additionally, some scholars have expressed their concern regarding the feedback function on preprint servers due to its potential to cultivate a toxic culture around preprints.[52] The function allows for the possibility of leaving unconstructive criticism or hostile comments on a preprint, especially considering that there is no gatekeeping and hardly any moderation on this function. Consequently, the reliability and quality of the feedback provided on preprints can be called into question. This is not to say that there is no value in the feedback that preprints receive, but rather that scholars need to remain critical.

This takes us to the second point of concern, that of the credibility and reliability of preprints. One of the main reasons why the credibility of preprints is questioned is due to the belief that, because the preprint lacks peer review, it must be of a lower quality than an article published in a peer-reviewed journal.[53] The content of a preprint has not been verified by experts in the field, nor has it been through rigorous peer review, thereby making it less

[51] Gross, 'Preprints and the Implications for Subsequent Peer Review and Publishing', p. 855.
[52] Puebla, Polka, & Rieger, *Preprints,* p. 27.
[53] Kodvanj, et al., 'Publishing of COVID-19 Preprints', p. 1340.

credible and reliable. While most preprint servers do have a basic screening process to establish that the paper reports scientific research by analyzing the structure and format, this is unlikely to vet out every misleading scientific manuscript that is posted on the server. Consequently, a reader has to be critical of a preprints credibility and reliability when reading and may, therefore, not trust the information as much as when it is presented through a peer-reviewed journal. This lack of trust in the credibility of preprints is a disadvantage to both readers and authors.

Furthermore, preprints could contain misleading or inaccurate information, so when these are cited, it could lead to misleading preprints being presented as verified information, which could damage the overall quality of scholarly knowledge.[54] This issue becomes even more prevalent when taking into account that there is a growing population of opportunists who post 'crackpot pseudo-science and misleading public health information' and present it as scholarly knowledge.[55] Moreover, due to preprints being openly accessible to everyone, these misleading preprints can be found by the general public, who are less prepared and skilled to assess the quality and credibility of scholarly work. Consequently, they may accept preprints as fact and may reinforce the spread of misinformation. Likewise, news outlets might do the same, and they can reach a much wider audience, allowing them to expose the public to unverified information.[56] Again, this does not imply that every preprint will spread misinformation, but rather that the public and media should remain critical and not present it as if the information has been certified. Media outlets often report on information found in preprints without mentioning this preprint status, which poses a significant risk, especially in fields such as medicine.[57] This also happened during the COVID-19 pandemic, which caused the spread of misinformation and lowered the overall public's trust in scientific knowledge.[58] To combat this, preprint servers added banners that display a disclaimer emphasizing that the material is a preprint and has not been peer-reviewed.[59] However, despite such a disclaimer, it is often either overlooked or ignored, meaning that these are not entirely effective. Therefore, while the accessibility of preprints is often considered an advantage of preprint, it does come with risks and can be classified as a disadvantage as well.

---

[54] Hoy, 'Rise of the Rxivs', 86.
[55] R. Anderson, 'Revisiting – Journalism, Preprint Servers, and the Truth: Allocating Accountability', *Scholarly Kitchen*, <https://scholarlykitchen.sspnet.org/2022/01/12/revisiting-journalism-preprint-servers-and-the-truth-allocating-accountability/> (13 June 2023).
[56] Penfold & Polka, 'Technical and Social Issues', p. 8.
[57] Ibid., p. 8.
[58] Puebla, Polka, & Rieger, *Preprints*, p. 17.
[59] Penfold & Polka, 'Technical and Social Issues', p. 8.

Another disadvantage is a common fear among scholars who are convinced that posting a preprint increases the risk of getting 'scooped', which lowers the trust in preprints even further. Getting 'scooped' refers to the situation in which a scholar posts their ideas or results as a preprint, which allows another researcher to 'scoop' this idea or result and publish it as their own intellectual property before the scholar who posted it as a preprint can do so.[60] Many scholars, especially those in fields where preprints are not commonly used, express concerns about their work being scooped if they post it as a preprint first. However, while it is understandable that scholars are cautious due to this reason, most papers addressing this risk also point out that the risk of getting scooped is minimal.[61] Puebla, Polka, and Rieger go as far as stating that there is no difference in risk of getting scooped in comparison to journal publication because preprint are time-stamped public records which can be used to claim priority.[62] Therefore, the fear of getting scooped, albeit being a collectively concern, is an ungrounded one. However, this concern does result in a lack of trust in some disciplines towards the use of preprints, which negatively impacts its status in scholarly communication. This is especially true in those disciplines who do not use preprints as much and thus have less experience with it in order to trust it.

A further disadvantage revolves around the question whether a scholar can get their work published after they have posted it as a preprint. While publishers have become more lenient about preprints and have updated their policies to include preprints, as mentioned before, several publishers have made it clear that they will not allow scholarly work to be published if it has been posted as a preprint.[63] Moreover, some journals are very specific with the kind of license a preprint should have if it is to be considered for publication.[64] There are no consistent guidelines on which license to choose as various publishers have differing policies on the subject. So, depending on the publisher and their policies, it might be disadvantageous for an author to post a preprint since it might prevent them from getting this work published in a journal later.

Lastly, there is a risk of preprints causing an 'overload' of information available to the public.[65] The vast amount of information available to the public is a known consequence of the digital age and the rise of the Internet, but with preprints this would also come to include

[60] Puebla, Polka, & Rieger, *Preprints,* p. 14.
[61] Teixeira da Silva, 'The Preprint Debate', p. 163; Puebla, Polka, & Rieger, *Preprints,* p. 14; and Penfold & Polka, 'Technical and Social Issues', p. 8.
[62] Puebla, Polka, & Rieger, *Preprints,* p. 14.
[63] Hoy, 'Rise of the Rxivs', 87.
[64] Puebla, Polka, & Rieger, *Preprints,* p. 17-18.
[65] Hoy, 'Rise of the Rxivs', 86.

more scholarly information. This would make finding the right information and sources much more difficult for the reader due to the vast amount of scholarly work made available. Furthermore, due to the fact that anyone can post information, the filter provided by academic publishers is omitted, which means that papers deemed irrelevant or uninteresting will also be posted as preprints.[66] Where the publisher usually fulfils the role of filtering for the reader, the reader is now tasked with this themselves, which is time-consuming and arduous.

Highlighting these disadvantages is not meant to diminish the potential that preprints have, nor is it meant to downplay the advantages. This subsection has merely aimed to provide a more well-rounded overview of preprints in general to inform the reader and ensure they have a solid understanding of the multifaceted nature of preprints before continuing on with this thesis.

1.5 Differences Per Discipline

It has already been alluded to throughout this chapter, but not every academic discipline has adopted preprints in the same degree. While preprints originated in the 1960s with biology, it was soon most popular within the field of physics. The first preprint server, arXiv, was launched in 1991 and where physics and related science fields embraced these servers, the life sciences remained hesitant. Despite efforts from both private and public sectors to establish preprint servers for biology, preprints have only become widespread within the life sciences in the last few years.[67]

Posting preprints has been a standard practice in the fields of mathematics, physics, economics and computer sciences, and in recent years it has also been adopted by many fields in the life sciences, such as biology, neuroscience, ecology and bioinformatics.[68] It is important to note that these fields fall under the broader category of STM (Science, Technology, and Medicine). However, it should be mentioned that the adoption of preprints in the humanities, social sciences, and arts (referred to as HSS) is relatively limited. Finding literature on the usage of preprints in HSS is challenging, as very few scholars discuss preprints in conjunction with HSS fields. When they are mentioned, it is usually within the context of psychology, which is technically classified as part of social sciences. Psychology

[66] Hoy, 'Rise of the Rxivs', p. 86-87.
[67] Penfold & Polka, 'Technical and Social Issues', p. 7.
[68] Ibid., p 7.

falls under the quantitative social sciences due to its empirical methods which is why it can overlap with STM disciplines, unlike the qualitative social sciences such as political studies.

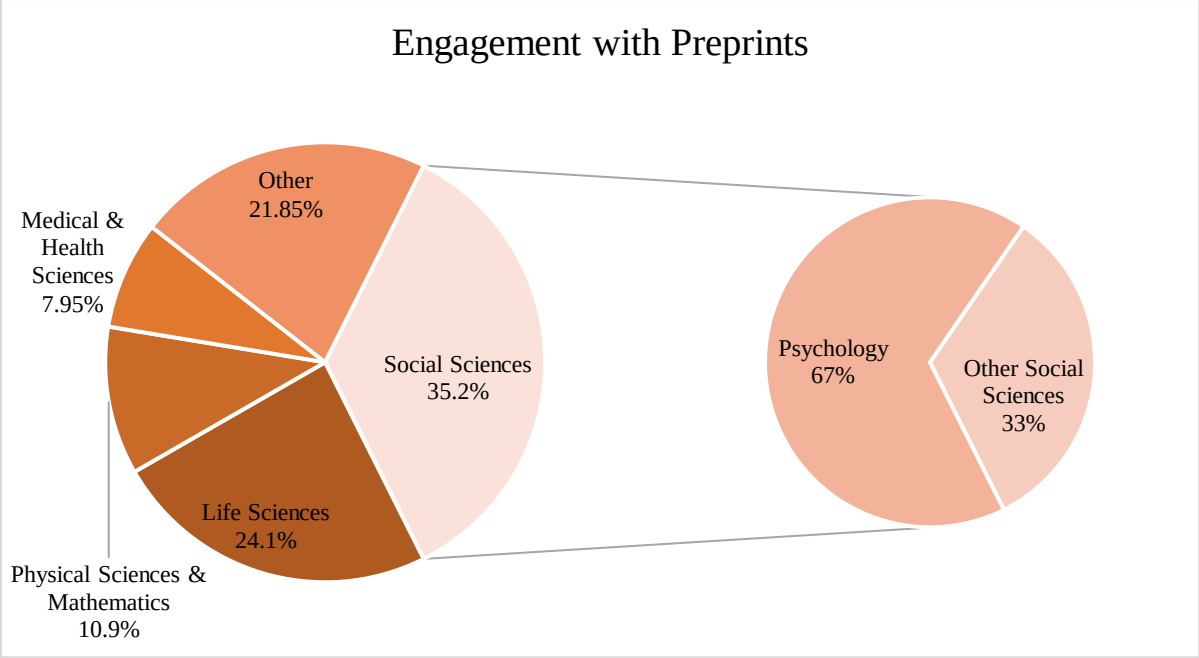

Figure 2: Engagement with Preprints Per Discipline According To a 2020 Survey.

A survey conducted in 2020, encompassing over 3500 researchers, identified four primary categories of disciplines that engage with preprints.[69] Social sciences, with psychologists comprising 67% of the respondents, accounted for 35,2% of the sample. Life sciences represented 24,1%, physical sciences and mathematics constituted 10,9%, and medical and health sciences accounted for 7,95%. It is worth noting that the inclusion of social sciences in these statistics primarily stems from the field of psychology, which overlaps with certain topics in the life sciences. The survey explicitly acknowledged that disciplines such as business, law, engineering, art, and humanities were not considered due to their negligible representation.[70] Therefore, these disciplines all fall under the factor 'other' in this graph, along with the researchers in the survey who did not signify their discipline.

Before further examining this divide between the acceptance of preprints in STM and HSS, this thesis will further elaborate on the increased usage of preprints in the field of psychology, while other social sciences have been hesitant to embrace it. The rationale behind

[69] C.K. Soderberg, Errignton, T.M., & Nosek, B.A., 'Credibility of Preprints: An Interdisciplinary Survey of Researchers', *Royal Society Open Science*, 7 (2020), pp. 1-17 (p. 4).
[70] Ibid., p. 11.

psychology's adoption of preprints mirrors the adoption seen in other life science disciplines following the popularity of preprints in medical fields during the COVID-19 pandemic. As highlighted by Puebla, Polka, and Rieger, the 'adoption of preprints by one life sciences community tends to result in their uptake by communities with overlapping interests'.[71] Due to interdisciplinary research, which overlapped with clinical research related to COVID-19, other life sciences fields became increasingly more familiar with preprints. As a result, researchers in other life sciences fields began adopting preprints as well. In contrast, when a field has minimal exposure to preprints, it lacks familiarity with both their benefits and drawbacks, which discourages their usage.[72] Thus, the reason why psychology, unlike other social science fields, has embraced preprints is because certain subfields of psychology align closely with STM fields such as biology and neuroscience. The implication here is that increased familiarity with and exposure to preprints encourages the adoption of preprints in their scientific fields. Consequently, one can infer that since psychology has started adopting preprints more frequently, this trend may eventually extend to other social sciences. However, only time will determine the validity of this hypothesis.

Furthermore, in earlier sections of this chapter, preprints have already been associated with Open Access, particularly green Open Access. This connection allows for a link to be made between the attitude towards Open Access in the HSS compared to STM, and their respective adoption of preprints. Given that preprints can be viewed as a form of green Open Access, it follows logically that disciplines exhibiting greater reluctance towards Open Access would also display more hesitation towards sharing preprints. Historically, STM disciplines have shown a significantly higher acceptance rate of Open Access models compared to HSS disciplines.[73] Consequently, it is reasonable to speculate that these contrasting attitudes towards Open Access may impact the popularity of preprints in these distinct fields to some extent.

Besides these disciplinary practice-based factors, various other factors have contributed to the differences in acceptance and usage of preprints across different disciplines, such as the importance or value of fast dissemination. While certain disciplines find fast dissemination through preprints highly advantageous, others consider it less important or attach less value to it. Most HSS fields place less importance on speed, while fields associated

[71] Puebla, Polka, & Rieger, *Preprints*, p. 36.
[72] Penfold & Polka, 'Technical and Social Issues', p. 7.
[73] J. Gross & Ryan, J.C., 'Landscapes of Research: Perceptions of Open Access (OA) Publishing in the Arts and Humanities', *Publications* 3 (2015), pp. 65-88 (p. 65).

with medicine or clinical research, for instance, attach less value to it. Instead they place great value on subjecting research to critical evaluation due to the potential risks faulty medical research can pose to public health. However, it is worth noting that the COVID-19 pandemic, with its high fatality rate, presented an exception to this trend. Typically, researchers in these fields prioritize the longer process of peer review over publication speed, resulting in less motivation to embrace preprints.[74]

Similarly, disciplines with shorter publication timelines or less pressure to rapidly publish tend to be less inclined to adopt preprints. For example, the field of chemistry has a much faster publication process compared to life sciences, which reduced the incentive to use preprints for their faster dissemination.[75] Many HHS fields are also unlikely to adopt preprints because there is less time-pressure for their work to be published. So, even though the publication timelines may not all be as fast as, for example, in the field of chemistry, there is still less of an incentive in these fields for the fast dissemination afforded by preprints.

It is likely that there are numerous other attributes and motivations that contribute to the difference between the disciplines in regard to preprints. However, this thesis does not aim to extensively delve into these disparities. The objective of this subsection was merely to acknowledge this difference and elaborate on possible explanations for these differences. The intention was to make the reader aware of the fact that preprints are not universally embraced across all academic disciplines and that in practice the situation is much more nuanced.

[74] Puebla, Polka, & Rieger, *Preprints*, p. 36.
[75] Ibid., p. 36.

Chapter 2. Theoretical Framework

In this chapter, I will analyse five theories addressing the functions of (academic) publishing in order to compare them and establish a comprehensive theoretical framework on which the subsequent analysis of preprints in the following chapter will be based. Each of these theories offers valuable insights that contribute to a theoretical framework that fits within the digital age and is focussed on academic publishing specifically. None of these five following theories fit this description well enough to allow for an analysis of preprints in relation to academic publishing. Some are focussed on publishing in general and others do not fit in the current digital age, which is a necessary condition for analysing preprints, therefore making these theories outdated. In order to formulate a theoretical framework on the functions of academic publishing in the current, digital age, this thesis will compare and combine five theories.

These five theories are drawn from Bhaskar's *The Content Machine,* Clark and Phillips' *Inside Book Publishing,* Origgi and Ramello's 'Current Dynamics of Scholarly Publishing', Roosendaal and Geurts' 'Forces and Functions in Scientific Communication', and Anderson's *Scholarly Communication.* The order in which these theories are mentioned here is also the order in which they will addressed in this chapter since it goes from most generic theory to a specific focus on academic publishing. Each of these theories has its own subsection in which they will be examined in depth. This chapter commences with Bhaskar's theory, which is the broadest theory among the five. It is followed by Clark and Philips' theory, which also identifies the functions of publishing in general, and then Origgi and Ramello, who provide an outsider's perspective on academic publishing. The theories of Roosendaal and Geurts and Anderson are addressed last and specifically focus on the functions of academic publishing.

After a critical comparative analysis of each of theories, this thesis will establish its own theoretical framework for the functions of academic publishing. This framework will allow for an analysis of preprints and their importance as a complementary tool for each function of academic publishing in the subsequent chapter.

2.1 M. Bhaskar's The Content Machine

As the subtitle suggests, in *The Content Machine: Towards a Theory of Publishing from the Printing Press to the Digital Network*, Bhaskar sets out to provide a universal theory on

publishing. Consequently, this means that his theory is not specifically about academic publishing, yet it is included in this thesis because analysing both the theories on publishing in general and the theories specifically focussing on academic publishing will result in a more comprehensive, nuanced framework on what the functions of academic publishing are. Additionally, Bhaskar makes it a point to address that his 'theory of publishing', whilst mainly focussed on book publishing, is applicable to all publishing contexts, meaning that it ranges from publishing a book to making a CD. He chose to do so because, according to him, there was no theoretical background with which to discuss publishing, which is why he wrote his theory in order to lay the foundation for a theory of publishing. Because of his broad definition of publishing, it also includes academic publishing among many other forms of publishing.[76]

In his theory, Bhaskar identifies four functions that are inherent to publishing and cannot be removed, unlike many other aspects.[77] These four core functions are framing, modelling, filtering and amplifying.[78] However, while Bhaskar identifies four functions, this thesis will only be analysing three of them, namely framing, filtering and amplifying. This is due to the fact that, according to Bhaskar, filtering, framing and amplifying are the 'how' of publishing, while modelling is the 'why'.[79] Albeit an important aspect of publishing as well, this thesis is not focussed on the 'why' of publishing since Bhaskar refers to business models and other financial models when discussing this 'why' and this aspect of publishing is less relevant for the subsequent analysis of preprints because preprints are freely available and do not generate a profit.[80] Rather, this thesis focusses on the process of publishing itself rather than the financial models behind it, so Bhaskar's modelling function has been excluded from this analysis.

In *The Content Machine,* Bhaskar stresses that content is a prerequisite for publishing, meaning that content is necessary in order to publish, thus, an understanding of the workings of content is essential in fully comprehending publishing. This is where *framing* comes in. As Bhaskar defines it, the frame is 'that which content *fills*'.[81] Bhaskar draws on the theorist Marshall McLuhan, who coined the phrase 'the medium is the message', arguing that it is the

[76] M. Bhaskar, *The Content Machine: Towards a Theory of Publishing from the Printing Press to the Digital Network* (London: Anthem, 2013), p. 133.
[77] Ibid., p. 103 and p. 133.
[78] Ibid., p. 6.
[79] Ibid., p. 137.
[80] Ibid., p. 97.
[81] Ibid., p. 79-80. (Original emphasis.)

medium which determines how the message is interpreted.[82] This means that the container of content, the frame, will shape the content itself, thereby affecting how it is perceived because each presentation of content is accompanied by certain preconceptions and expectations.[83] With the shift to digital environments, new frames have been created, such as codes, servers and screens.[84] These new frames come with their own sets of preconceptions that will influence how the content is perceived. Bhaskar often emphasizes how publishing is about constant change. While the functions of publishing remain constant at their core, the way in which these functions are fulfilled changes with time. Framing is continually changing, not only through the invention of new frames, but also through constantly changing expectations that audiences have of these frames, which is something publishers have to be critically aware of.

Another function of publishing which Bhaskar identifies in *The Content Machine* is *filtering*. Filtering plays an integral part in publishing as it involves the selection of what content publishers will frame and what content they will not be framing. Interestingly, instead of using the term 'selection' to describe this function of publishing, Bhaskar deliberately employs the term filtering, due to its inclusivity and its implication of a broader range of possibilities within the realm of publishing.[85] By using the term 'filtering', Bhaskar emphasizes the increasingly active role of publishers in curating content, particularly in the digital era where an abundance of content is readily available. Unlike the term 'selection', filtering implies that publishers are allowed to exercise their agency in carefully choosing and presenting content. Bhaskar asserts that '[f]iltering is slowly shifting from a selection emphasis to a curation emphasis'.[86] This is relevant within the context of academic publishing as well since academic publishers receive an abundance of manuscripts and their objective is to filter out the most relevant and innovative works, meaning that their focus is more on curating than on selecting. Through his use of terminology, Bhaskar aims to establish a more inclusive and comprehensive understanding of this function of publishing that is supposed to be universal and thus transcends time periods.

Lastly, Bhaskar addresses another function that cannot be removed from publishing, namely *amplifying.* Amplification is 'at the heart of publishing' and Bhaskar defines it as

[82] M. McLuhan, *Understanding Media: The Extensions of Man* (New York: MIT Press, 1994), p. 7.
[83] Bhaskar, *The Content Machine*, p. 84.
[84] M. Bhaskar, 'Filtering, Framing, Amplifying: The Core of Publishing Now and Then', *TXT*, 1 (2014), pp. 78-81 (p. 80).
[85] Bhaskar, *The Content Machine*, p. 107.
[86] Ibid., p. 185.

'ensuring that a work is more widely encountered than without the amplifying act. That's it; that's publishing.'[87] Bhaskar argues that making something public is not the same as publishing, since, according to his own example, that would make leaving a manuscript on a park bench – thereby 'making it public' – the same as publishing. To address this issue, Bhaskar proposes the term 'amplification' as a more concrete and accurate description of the publishing process. He highlights the importance of framing in amplification because publishers use framing techniques to enhance the reach and impact of a work.[88] Without an appropriate frame, there would be nothing to amplify. The right frame, on the other hand, can greatly benefit the effect of amplification. With the advent of the digital age, amplification has undergone significant changes. Instead of solely focusing on producing and distributing works to make them available, amplification now revolves around drawing attention to a particular work.[89] For publishers, the challenge lies in capturing the attention of the audience amidst the abundance of content available online. As technology continues to shape the way information is consumed, publishers must adapt their amplification strategies to ensure that works are not only made public but also encountered by a broader audience.

While Bhaskar's theory is too generic for this thesis' analysis of preprints, this wide scope is also useful as a starting point from which to build the theoretical framework that is applicable to preprints and academic publishing. Bhaskar puts an emphasis on establishing a theory that covers publishing in a way that transcends time periods. Its broad scope makes it applicable to both publishing from centuries ago and publishing in the digital age, which is what makes this theory valuable.

2.2 G. Clark and A. Phillips' Inside Book Publishing

Similarly to Bhaskar's *The Content Machine,* Clark and Phillips' *Inside Book Publishing* has a more general focus on publishing, rather than specifically concentrating on academic publishing. Nevertheless, like Bhaskar's theory, this work by Clark and Phillips is also included in this chapter for the same reason. By incorporating theories encompassing publishing in general and theories that specifically address academic publishing, a more comprehensive framework can be established due to it containing multiple perspectives.

[87] Bhaskar, 'Filtering, Framing, Amplifying', p. 80.
[88] Bhaskar, *The Content Machine*, p. 115.
[89] Bhaskar, *The Content Machine*, p. 177; Bhaskar, 'Filtering, Framing, Amplifying', p. 81.

Therefore, this thesis will also be analysing the functions of publishing that Clark and Phillips put forth.

These elements that Clark and Phillips have identified can be subdivided into three categories or, in other words, three primary functions. The numerous elements that they specify in *Inside Book Publishing* are 'the curation and acquisition of intellectual property, editorial, design and production, marketing, and sales'.[90] According to Clark and Phillips, these are the elements through which publishers add value to an author's work.[91] They also acknowledge that these elements can be combined in separate functions. For instance, editorial, design, and production can be amalgamated into the function of product development and investment.[92] Similarly, the curation and acquisition of intellectual property can be merged into the function of selection, and the elements of marketing and sales can be combined in the function of connecting markets. To summarize, Clark and Phillips identify the three primary functions of publishing as selection, product development and investment, and connecting markets.

First, the *selection* function of publishing encompasses both curation and acquisition processes, which makes it comparable to Bhaskar's function of filtering.[93] In regard to curation, Clark and Phillips emphasize that the publisher's role is that of 'the curator of excellent content'.[94] With this curation of content, publishers endorse selected works, granting them a valuable brand recognition. By lending its brand, a publisher signifies that a work is worthy of publication.[95] The brand endorsement of an academic publisher provides authors with recognition from their peers and opportunities to advance their careers.[96] In regard to acquisition, when a publisher selects a work they find worthwhile to publish, they obtain 'an exclusive licence to exploit the intellectual property rights'.[97] This grants them the authority to publish the author's work in both print and electronic forms. Simultaneously, it becomes the publisher's responsibility to safeguard the author's rights against infringement by others, utilizing technical and legal measures.[98] In the digital age, where file sharing has become increasingly prevalent, protecting these rights has become crucial. Together, curation and acquisition comprise the selection function of publishing. By fulfilling these roles, publishers

[90] G. Clark & A. Phillips, *Inside Book Publishing* (New York: Routledge, 1988 (6th ed., 2020)), p. 116.
[91] Clark & Phillips, *Inside Book Publishing*, p. 116.
[92] Ibid., p. 116.
[93] Bhaskar, *The Content Machine*, p. 107.
[94] Clark & Phillips, *Inside Book Publishing*, p. 117.
[95] Ibid., p. 117.
[96] Ibid., p. 117.
[97] Ibid., p. 119.
[98] Ibid., p. 119.

contribute to the dissemination of high-quality content while ensuring the protection and recognition of authors' intellectual property.

Next, there is the publishing function of *product development and investment*, which entails editorial, design, and production services. Once an author's work is selected, one of the primary services provided by a publisher is their editorial expertise. As highlighted by Clark and Phillips, this editorial process can encompass a broad spectrum, ranging from providing structural advice to line-by-line editing.[99] The objective of such editing is to ensure that the author's work is coherent and easily comprehensible to readers. Additionally, the design and production values of a published work also play a vital role. Clark and Phillips identify several key aspects to consider, including the length, format, and cover design. Moreover, they emphasize the significance of ensuring the quality, accuracy, and relevance of the content.[100] Although some of these tasks may not be immediately associated with academic publishing, even the design of a cover can be of importance when it comes to publishing journals or academic books, such as monographs. Academic publishers make substantial investments in the production of the works they acquire. In the digital era, these publishers have had to adapt to the advent of online distribution channels and various devices.[101] This shift has necessitated their facilitation of a seamless transition to the digital environment. When combined, the values of editing, design, and production form the core of the function of product development and investment in publishing. Each aspect contributes to enhancing the overall quality of the published work, ensuring its appeal to the target audience.

Finally, Clark and Phillips have identified the key elements of marketing and sales, which jointly serve as the function of *connecting markets*. The primary goal of publishers is to market a work in a manner that optimizes sales and generates the highest possible profit.[102] While monetary gain may not be a relevant factor for authors in academic publishing, achieving high sales volumes can signify a larger readership, leading to increased visibility for their work. This aligns with the concept of branding mentioned earlier. By choosing an academic publisher that also strives to maximize sales, thus also increasing the readership, an author can enhance the visibility of their work among their peers, which may advance career opportunities. This is where it becomes evident that Clark and Phillips were speaking of publishing in general because their function of connecting markets does not account for Open

[99] Clark & Phillips, *Inside Book Publishing*, p. 118.
[100] Ibid., p. 118.
[101] Ibid., p. 118.
[102] Ibid., p. 119.

Access journals. This thesis will elaborate on this issue specifically further on in this chapter. Aside from maximizing sales, publishers employ strategies like search engine optimization to maximize the discoverability of an author's work, making it easier for the general public to find.[103] Furthermore, in terms of sales, larger publishers have a stronger position to negotiate the distribution of an author's work with key retailers worldwide, such as libraries.[104] Having one's work available through a library further increases the discoverability and readership, which, as mentioned above, provides validation for the author. Taken together, these marketing and sales elements, although seemingly less crucial in academic publishing, are still essential for academic authors.

Of the three primary functions Clark and Phillips identify, only the function of selection appears to strongly overlap with Bhaskar's filtering function.[105] There is some similarity between Bhaskar's amplifying function and Clark and Phillips' connecting markets as well, although these similarities are less prominent. The two functions have different objectives, Bhaskar's being about a wider dissemination and Clark and Phillips' focussing on optimizing financial gain.[106] However, these functions have the same result, namely gathering a larger readership. Overall, this theory, like Bhaskar's, is too broad to analyse preprints in relation to academic publishing, but such a broad perspective can be useful when building my own framework.

2.3 G. Origgi and G. Ramello's 'Current Dynamics of Scholarly Publishing'

Origgi and Ramello's paper, 'Current Dynamics of Scholarly Publishing', published in 2015, introduces a research project focussed on elucidating the current dynamics of academic publishing.[107] As mentioned in the introduction of this thesis, Origgi and Ramello are outsiders to the field of publishing, having a background in philosophy and industrial economics respectively. With their outsiders' perspectives, they can bring new insights to the theory of academic publishing, which might differ from the perspectives belonging to scholars who are heavily involved in the publishing industry and can therefore be a valuable addition to the nuanced framework that this thesis will formulate.

[103] Clark & Phillips, *Inside Book Publishing*, p. 119.
[104] Ibid., p. 116 and 119.
[105] Bhaskar, *The Content Machine*, p. 107.
[106] Bhaskar, 'Filtering, Framing, Amplifying', p. 80.
[107] G. Origgi & G.B. Ramello, 'Current Dynamics of Scholarly Publishing', *Evaluation Review*, 39 (2015), pp. 3-18 (p. 4).

Origgi and Ramello's goal is to establish a comprehensive framework that encompasses the ongoing debates surrounding this field.[108] To accomplish their goal, they examine the main components that characterize academic publishing, thereby identifying its primary functions. By elucidating these functions, Origgi and Ramello offer a comprehensive understanding of academic publishing. Their research project aims to provide valuable insights into the dynamics and complexities of this domain, which contribute to shaping the future of academic publishing. The functions they identify are dissemination of knowledge, expansion of scholarly debate, and validation of scholarship.

The main objective of academic publishing, which Origgi and Ramello specify, is the *dissemination of knowledge.*[109] This function of academic publishing highlights the significance of journal publications in effectively spreading scientific advances to the wider academic community, which is comparable to Bhaskar's amplifying function.[110] The primary function of academic publishing is to provide a platform for researchers to share their findings and insights with others in their respective disciplines. By publishing their work in reputable journals, researchers make their discoveries accessible to a global audience of peers, which enables the scientific community to build upon existing knowledge and make significant progress. As Origgi and Ramello indicate, the creation of scholarly journals stems from the goal of disseminating knowledge, ensuring that scholarly information becomes accessible to the public.[111] This accessibility is instrumental in bridging the gap between researchers and the wider audience, allowing individuals to engage with scholarly information they might otherwise not have encountered. Consequently, Origgi and Ramello's emphasis on the dissemination of knowledge as a primary function of academic publishing is well-founded.

Another key function of academic publishing, according to Origgi and Ramello, is the *expansion of scholarly debate*, which builds upon the function of disseminating knowledge.[112] As mentioned above, academic publishing functions as a platform for researchers to share scholarly information with a wider audience. Therefore, in addition to wide disseminating knowledge, it also broadens the possibility for scholarly debate. It allows more researchers to respond to, challenge, or improve existing theories and discoveries. This dynamic interaction leads to further insights and innovation. Moreover, academic publishing actively encourages scholarly discourse and intellectual exchange. During the publication process, scholars

[108] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 5.
[109] Ibid., p. 8.
[110] Bhaskar, 'Filtering, Framing, Amplifying', p. 80.
[111] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 8.
[112] Ibid., p. 8.

participate in peer review, a rigorous evaluation conducted by experts in the relevant field. This critical assessment ensures the quality and credibility of the research, while also providing valuable feedback and suggestions.[113] Such a process stimulates scholarly debate, prompting researchers to address criticisms and refine their arguments, ultimately contributing to a more robust and nuanced understanding of the subject matter. As Origgi and Ramello assert, scientific journals were established mainly with this goal to 'expand the boundaries of scholarly debate much more widely than could be done in person'.[114] Therefore, the expansion of scholarly debate is one of the primary functions of academic publishing.

Lastly, Origgi and Ramello identify the *validation of scholarship* as another primary function of academic publishing.[115] As mentioned previously, scholarly works go through a certain process of quality control, predominantly peer review, which encourages critical engagement with the topic while also ensuring that the content is not only relevant but also correct. Experts in the relevant field will go over a scholarly work and assess whether the research it presents is valid. This process ensures that any work published has gone through a rigorous evaluation of its content and credibility. This inadvertently means that the author's work receives a stamp of approval from the publisher and, thereby, validates their scholarly work. Furthermore, Origgi and Ramello assert that it is through the critical selection and validation of scholarly works that academic publishers encourage the competition between scholars, which results in a selection of the highest quality of scholarly works.[116] Of all the submission an academic publisher receives, only a fraction ends up being published in their journals, so scholars face intense competition in their quest to have their work published. Consequently, they exert great effort to produce scholarly works of the utmost quality, in the anticipation that their submissions will be accepted and published by reputable academic publishers. In short, the validation of scholarship is a crucial function of academic publishers through which they contribute to the advancement of knowledge and the preservation of academic excellence.

In the theory that Origgi and Ramello present, the two functions of dissemination of knowledge and expansion of scholarly debate do seem to overlap in some aspects. It appears that dissemination is the act of the publisher and expansion of debate is one of the

[113] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 4.
[114] Ibid., p. 8.
[115] Ibid., p. 8.
[116] Ibid., p. 8.

consequences of wide dissemination. Additionally, Origgi and Ramello's dissemination function is comparable to the amplifying function that Bhaskar identifies, both addressing the act of disseminating a work to a larger audience than it otherwise would have encountered.[117] Moreover, a valuable addition from Origgi and Ramello is the identification of validation as a separate function of academic publishing. In Clark and Phillips' theory they do address how brand endorsement from a publisher is a form of validation for the author, but in their theory this was an aspect of the selection function, rather than a separate function on its own.[118] Origgi and Ramello's paper puts emphasis on the significance of validation by classing it as a primary function of academic publishing, rather than a smaller aspect of one of the functions of publishing.

2.4 H. Roosendaal and P. Geurts' 'Forces and Functions in Scientific Communication'

Roosendaal and Geurts presented this paper, 'Forces and Functions in Scientific Communication: An Analysis of Their Interplay', at a conference in Germany in 1997. This paper presents an examination of 'the transformation of the linear scientific information chain into a scientific communication "network"'.[119] The authors aim to establish a methodology for studying this transformation. Within the framework of their developed methodology, they have identified four key functions of scientific communication. These four functions are registration, awareness, certification, and archiving.

Roosendaal and Geurts extend the applicability of their findings by recognizing that the functions they identify are also relevant to the realm of academic publishing. They support their assertion by showcasing how the four communication functions can be exemplified within the context of the first research journal.[120] It is on this account that Roosendaal and Geurts' theory regarding the functions of scientific communication is included in this thesis to establish a framework about the functions of academic publishing.

The aforementioned functions that Roosendaal and Geurts identify – registration, awareness, certification, and archiving – can be divided in two separate axes. In order to clarify this, Roosendaal and Geurts have created the following illustration:

[117] Bhaskar, 'Filtering, Framing, Amplifying', p. 80.
[118] Clark & Phillips, *Inside Book Publishing*, p. 117.
[119] H.E. Roosendaal & P.A.T.M. Geurts, 'Forces and Functions in Scientific Communication: An Analysis of Their Interplay', presented at Conference on 'Cooperative Research in Infortmation Systems in Physics', 1997, pp. 1-32 (p. 26).
[120] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 15-16.

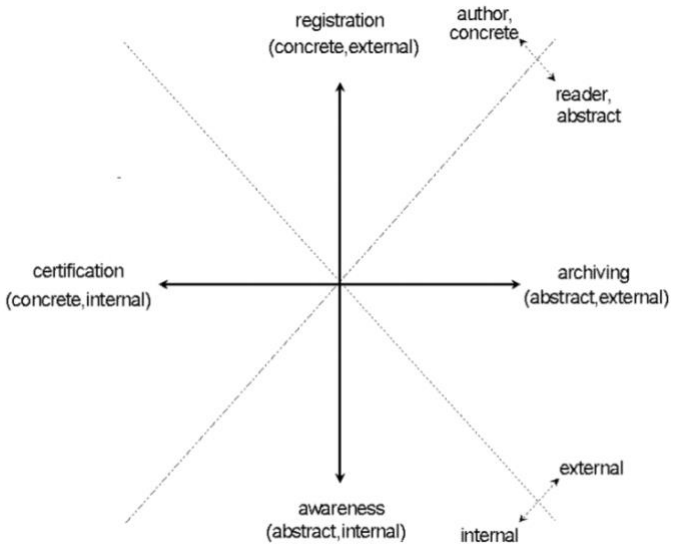

Figure 2: The Four Functions of Scientific Communication.[121]

The horizontal axis depicts the different aspects of scientific judgement, certification and archiving, while the vertical axis depicts the different aspects of scientific observation, registration and awareness.[122] The diagonal lines behind these axes are the four main market forces in scientific communication, which Roosendaal and Geurts establish earlier on in their paper. This thesis does not aim to extensively explore the intricacies of these market forces since these are irrelevant to this thesis. It will only touch upon them when essential to clarify the four main functions that are presently the central focus of this thesis.

One of the primary functions in scientific communication is *registration*.[123] The significance of registration lies in its ability to establish priority, ensuring that the rightful individual receives credit for their scientific findings, which safeguards against any potential disputes or claims of precedence from other parties. [124] A widely employed method of registering new findings is by submitting them to a publisher, effectively making registration an integral function of academic publishing. Researchers often choose to submit their work to reputable publishers to ensure that their findings are formally recorded and made accessible to the scientific community. By doing so, they create a tangible record of their contributions, establishing their priority in the field and allowing others to build upon their work. In essence,

[121] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 14.
[122] Ibid., p. 14-15.
[123] Ibid., p. 15.
[124] Ibid., p. 17.

registration serves as a mechanism to validate and recognize scientific accomplishments, therefore, this can be connected back to Origgi and Ramello's validation function.[125] In regard to the registration function, academic publishers act as a crucial conduit in the registration and validation process by keeping a record of these registrations.

Following registration is the function of *certification,* which has been a widely contested function, with Roosendaal and Geurts theorizing that it is likely to undergo substantial development in the near future.[126] The peer review component of certification, in particular, has attracted considerable attention, which again illustrates that this function of scientific communication is also linked to academic publishing. Peer review has become a contentious issue, raising questions about its effectiveness and value as scholars have expressed doubts regarding the merits of this practice.[127] However, despite the debates surrounding it, some form of review remains crucial during the certification process. It is during this phase that revisions are undertaken to refine the work before it is deemed ready for publication. The future of certification, as Roosendaal and Geurts speculate in 1997, holds promising possibilities for further advancements. As the scientific landscape evolves, there is a growing need to reassess and enhance the certification function. This includes reevaluating the peer review system and exploring alternative methods that ensure the quality and credibility of scientific research.[128] However, where Roosendaal and Geurts claimed in 1997 that it would have experienced substantial change it the near future, we can now reflect back and determine that not much has changed in the peer review process since their statement. The main difference is that peer review has been adapted to digital workflows, like most processes in academic publishing, but it has not changed conceptually like Roosendaal and Geurts predicted.

Next, the function of *archiving* is an integral part of scientific communication, which can also be considered a primary function of academic publishing.[129] Roosendaal and Geurts consider archiving an external function since it is more easily outsourced, for example, to publishers.[130] Academic publishers have established digitized archives for scientific information and shoulder the responsibility of maintaining and preserving the archived information. They have taken initiatives to create digital repositories that house vast amounts

[125] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 8.
[126] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 15.
[127] T. Crick, et al., 'A Multi-Disciplinary Perspective on Emergent and Future Innovations in Peer Review', *F1000Research,* 6 (2017), pp. 1151-1215 (p. 1152).
[128] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 18.
[129] Ibid., p. 15.
[130] Ibid., p. 15.

of scientific data. These archives serve as valuable resources for researchers, enabling them to access and retrieve past findings, studies, and scholarly works. The responsibility of academic publishers extends beyond the mere storage of information. They must ensure the integrity and authenticity of archived content, maintaining accurate metadata and facilitating proper citation practices. Academic publishers invest in robust infrastructure and employ preservation strategies to protect the archived data from technological obsolescence and ensure its long-term accessibility.

Lastly, Roosendaal and Geurts have identified the *awareness* function of scientific communication, which they consider the most difficult function.[131] However, in their paper, Roosendaal and Geurts provide limited insight into this function. They neither provide a clear definition nor delve into it extensively. It is only when they discuss the example of the first journal that they vaguely hint at what awareness entails. They acknowledge the historical issue of uneven dissemination of scientific information to readers, which proved to be both ineffective and inefficient. The academic journal emerged as a solution to this issue.[132] Consequently, we can infer that the awareness function ensures that every member of the scientific community is informed about new publications, providing them with equal opportunities to discover them. This function becomes a crucial responsibility of academic publishers, who are tasked with ensuring public awareness of the articles and books they release.

In a way, this awareness function can be connected back to Origgi and Ramello's dissemination function and Bhaskar's amplifying function.[133] By disseminating a scholarly work to a wider audience, academic publishers raise awareness to that work. Like Origgi and Ramello's function regarding the expansion of scholarly debate, Roosendaal and Geurts' function of awareness is a consequence of wider dissemination. Yet, raising awareness to a certain work can also be linked to Clark and Phillips' connecting markets since the marketing aspect of this function also aligns with what Roosendaal and Geurts appear to describe.[134] Again, their paper does not elaborate much on the awareness function that they identify, it only provides an example with an indication of what they mean. Furthermore, a combination of Roosendaal and Geurts' registration and certification functions align with Origgi and Ramello's validation function, since both contribute to affirming the credibility and quality of

[131] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 15 and 18.
[132] Ibid., p. 16.
[133] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 8; and Bhaskar, 'Filtering, Framing, Amplifying', p. 80.
[134] Clark & Phillips, *Inside Book Publishing*, p. 119.

a work, thereby providing validation for the author when they publish with an academic publisher.[135]

2.5 R. Anderson's Scholarly Communication

Lastly, this chapter discusses Anderson's work, *Scholarly Communication: What Everyone Needs to Know*, which addresses some of the most important characteristics and developments of the scholarly communication system, as well as shedding light on its most controversial issues. Among the subjects explored in his work is academic publishing, which he considers as a subset of scholarly communication.[136] One chapter is dedicated to elucidating the nature of academic publishers, highlighting four crucial categories of service that academic publishers have historically provided. In other words, Anderson has identified four primary functions of academic publishing, namely selection, editorial services, making available, and branding and marketing.[137] Anderson observes that these functions have undergone changes throughout time and are still evolving, a subject he also examines in his analysis.

First, Anderson has identified the *selectivity function*, which entails the selection of submitted materials for quality and relevance.[138] This function is comparable to Bhaskar's filtering and Clark and Phillips' selection function.[139] Publishers have to critically select those articles and manuscripts that they deem most relevant and of the highest quality, and the core of this selectivity function has not changed much in the digital age. The selection process means that publishers have prioritized a specific article or manuscript, thereby sparing the audience the arduous task of having to individually assess numerous articles or manuscripts in their search for relevant and trustworthy information.[140] Nonetheless, Anderson cautions that despite the rigorous selection process undertaken by academic publishers, readers should still approach their reading material with a critical mindset.[141]

Second, the *editorial function* is another vital aspect of academic publishers, which entails that they provide 'editorial services and thereby [work] to improve and refine the

[135] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 8.
[136] Anderson, *Scholarly Communication*, p. 59.
[137] Ibid., p. 59.
[138] Ibid., p. 59.
[139] Bhaskar, *The Content Machine*, p. 107; and Clark & Phillips, *Inside Book Publishing*, p. 117.
[140] Anderson, *Scholarly Communication*, p. 60.
[141] Ibid., p. 61.

author's work'.[142] According to Anderson's observation, the manuscript of an article or book can contain valuable or innovative information, and yet its organization and writing style may unnecessarily be lacking in clarity.[143] The publisher's editorial services aim to enhance the readability and coherence of the original manuscript version, benefiting the authors, particularly those from lower or middle income countries where English may not be their first language and fluency in it might be limited.[144] In general, but especially in this aspect, the editorial function proves to be highly important and valuable. Additionally, these editorial services are employed to ensure that the text and citation formatting are adjusted to align with the publisher's established 'house style'.[145] Academic publishers have their own 'house style' in order to portray a certain consistency between the articles and books they publish. This function bears similarities to Clark and Phillips' function of product development because it not only ensures the scholarly work's quality but also its eventual presentation.[146] As Anderson suggests, the selection and editing processes have always been crucial in academic publishing.[147] Despite the profound changes that digital environments have brought about over the past few decades, altering the production and distribution of information to the public, these functions remain fundamentally important.

Third, Anderson discusses *making available* as a primary function of academic publishing, which is similar to Bhaskar's amplifying function and Origgi and Ramello's dissemination.[148] This function, unlike those of selecting and editing, has been significantly impacted by the changing digital landscape.[149] Formerly, it was challenging to uncover the existence of particular documents and this was usually reserved for a very privileged audience that had access to catalogues.[150] However, relying solely on catalogues had its drawbacks, as they were often outdated and rather limiting. The recent rise of the web has significantly simplified this task. The emergence of search engines has revolutionized scholarship by enhancing the accessibility and widespread availability of scholarly information. Still, Anderson raises a valid argument that being aware of the existence of a document does not equate to having the ability to access it. [151] In the past, during the era of print media, acquiring

[142] Anderson, *Scholarly Communication*, p. 59
[143] Ibid., p. 61.
[144] Ibid., p. 61.
[145] Ibid., p. 61.
[146] Clark & Phillips, *Inside Book Publishing*, p. 118.
[147] Anderson, *Scholarly Communication*, p. 62.
[148] Bhaskar, 'Filtering, Framing, Amplifying', p. 80; and Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 8.
[149] Anderson, *Scholarly Communication*, p. 62.
[150] Ibid., p. 62.
[151] Ibid., p. 62.

a copy of a document was sufficient to gain access, enabling one to easily share it with a friend. In today's digital age, this aspect of publishing has been significantly weakened since making intellectual property accessible to others has become easier, regardless of the value that publishers traditionally offered through their selectivity, editorial review, and quality control processes. Anderson concludes that 'the value publishers offer as distributors of content has been significantly undermined'.[152] Yet, he still identifies it as a primary function of academic publishing, meaning that he still attaches value to this function and he does not consider it to be nullified.

The last primary function of academic publishing, according to Anderson, is that of *branding and marketing*, meaning that a publisher endorses a document, which bestows a sense of prestige upon it and assists in garnering the attention of the author's peers, and possibly even the general public.[153] Due to the functions of selection and editorial services that publishers provide, any article or book published by an academic publisher has gone through critical scrutiny and is expected to be of a certain quality, therefore it gains prestige. In turn, this endorsement enhances the credibility and reputation of the author within the academic community. In this aspect, this function of Anderson can be compared to Origgi and Ramello's validation function.[154] Moreover, publishers actively engage in promoting an author's work and possess connections to various entities such as university libraries.[155] This aspect, on the other hand, is more similar to Clark and Phillips' connecting markets function.[156] This connection to scholarly institutions and resources enhances the value of publishing with an academic publisher for an author. It provides them with access to a wider audience and facilitates the dissemination of their ideas among fellow academics and researchers.

Each of the functions that Anderson identifies has overlap with the other aforementioned theories, demonstrating that this is an inclusive theoretical framework on the functions of academic publishing. The reason why this theory has not been used to analyse preprints is due to Anderson simplifying his theory to make it accessible for everyone, even those without knowledge about scholarly communication. This helpful approach is valuable for those learning about scholarly communication and academic publishing, but it does imply that some subtleties and deeper facets of these functions may be overlooked. By drawing on

[152] Anderson, *Scholarly Communication*, p. 63.
[153] Ibid., p. 59.
[154] Origgi & Ramello, 'Current Dynamics of Scholarly Publishing', p. 8.
[155] Ibid., p. 62 and 249.
[156] Clark & Phillips, *Inside Book Publishing*, p. 119.

the other four theories in addition to Anderson's theory, this thesis can formulate a more nuanced, comprehensive framework.

2.6 Comparative Analysis: Forming a Comprehensive Framework

Each of the above theories has put forth a number of functions of (academic) publishing, yet as I have already addressed throughout this chapter, many of these functions overlap. Despite phrasing these functions differently and bringing some varying nuances to them, many of these functions refer to the same thing and can thus be grouped together. Therefore, a comparison of these theories puts forth five functions of academic publishing that encompass and combine most of the functions that these scholars have addressed. These five functions are filtering, editing, validation, dissemination, and marketing.

This new framework incorporates nearly all the functions identified by these scholars, with the exception of archiving, proposed by Roosendaal and Geurts. Archiving might be an aspect of publishing, but it does not appear to be a primary function, and it remains unclear to what extent academic publishers actually fulfil this role. As Roosendaal and Geurts point out, archiving is an external function that can easily be outsourced and is also often handled by libraries as well.[157] Hence, archiving appears to be primarily associated with libraries rather than academic publishers, although publishers contribute to the process. Accordingly, archiving is excluded from this theoretical framework of the principal functions of academic publishing. The aforementioned five functions that it does consists of are filtering, editing, validation, dissemination, and marketing.

For starters, the function of *filtering* is referenced in the theories of Bhaskar, Anderson, and Clark and Phillips. This thesis has opted to use the term 'filtering' because, as Bhaskar argues, filtering is a more comprehensive term since it encompasses both selection and curation of scholarly works.[158] Filtering involves carefully assessing and selecting scholarly materials that meet the standards and objectives set by the publisher. Through this process, publishers make informed decisions about which works to publish, considering factors such as quality and relevance. By employing the term filtering to describe this function, this thesis recognizes the relevance of both the initial act of selecting academic work

[157] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 15.
[158] Bhaskar, *The Content Machine*, p. 107 and 185.

and the subsequent curation of this work to determine its value and significance to academic scholarship.

Another function of academic publishing is *editing*, which, despite it often being discussed or referenced in these theories, only Anderson and Clark and Phillips consider to be a primary function of publishing. In this thesis's framework, the editing function combines the aspects which Anderson refers to with those that Clark and Phillips address. Anderson's definition of the editorial function mainly focusses on lending services to help improve the writing style and overall readability, as well as ensuring that a paper is written in the 'house style' of a publisher.[159] In a way, this is comparable to Bhaskar's framing since it focusses on how content is presented, which will influence how it is perceived by an audience.[160] Then, where Clark and Phillips reference the editorial services as well, they also address the significance of ensuring the relevance, accuracy and quality of the content.[161] In order to achieve this objective, academic publishers make use of peer review. Clark and Phillips do not mention peer review explicitly, yet it does contribute to this editing function by ensuring that an academic paper is reviewed and critically evaluated. So, in addition to the smaller editorial services that help improve the writing style and readability, this thesis' editing function also includes peer review, which ensures a certain quality and accuracy.

As a result of both filtering and editing, academic publishers provide *validation*, a function explicitly referenced in the works of Anderson, Roosendaal and Geurts, and Origgi and Ramello. Each of these scholars roughly define it as an endorsement on the publisher's behalf for a scholarly work. The functions of filtering and editing, which include curation and peer review, mean that a publisher has validated the findings presented in the paper and has ensured their relevance. This function not only validates the scholarly work, but also the author who managed to publish their work with a certain academic publisher. Consequently, this function also somewhat encompasses what Roosendaal and Geurts refer to as 'registration'.[162] When an author's work is validated, it becomes recognized and when the author is the first to make a certain discovery, this validation automatically leads to the ability to claim priority. In short, the validation of a scholarly work is a primary function of academic publishing that also encompasses establishing priority.

[159] Anderson, *Scholarly Communication,* p. 59 and 61.
[160] Bhaskar, *The Content Machine,* p. 84.
[161] Clark and Phillips, *Inside Book Publishing,* p. 118.
[162] Roosendaal & Geurts, 'Forces and Functions in Scientific Communication', p. 15.

Then, there is the function of *dissemination*, which Bhaskar, Anderson, Roosendaal and Geurts, and Origgi and Ramello all identify. Dissemination refers to making scholarly work more widely available to a larger public and spreading it effectively to a wider academic readership. It means, as Bhaskar puts it, ensuring that a work is encountered by a wider audience than without the effort to amplify it.[163] As part of disseminating scholarly work, publishers ensure it is easy to find through search engines, raise awareness to its existence and make it available. All these aspects, which have been described by the aforementioned scholars, fall under what this thesis identifies as dissemination and is a primary function of academic publishing.

Finally, this thesis addresses one last function of academic publishing, namely that of *marketing*, which both Anderson and Clark and Phillips address. While marketing does tie into the function dissemination there is enough of a difference between the two that they have been identified as two separate functions. For this thesis I have considered grouping the two functions together, but on the grounds of certain nuances I opted to keep them separate. The function of marketing refers to an academic publisher's objective to reach their target audiences and convince them to read and engage with their academic journals. As a result, they aim to acquire, retain and expand this audience. One of the desired consequences of this is generating a revenue through licenses and subscription fees. Another consequence is the increase of brand awareness, which may draw in more academic authors who will submit their work because of the reputation these publishers have built. While generating a revenue does not occur in the case of Open Access journals, these journals do help increase brand awareness for the publishers, which is why the function of marketing still applies and is considered a primary function of academic publishing.

In short, through a comparative analysis of five theories addressing the functions of (academic) publishing, this thesis has compiled a theoretical framework consisting of five primary functions of academic publishing, namely filtering, editing, validation, dissemination, and marketing. These functions often relate or intersect with one another. Filtering and editing result in validation while dissemination and marketing have some overlap despite different motivations on the publisher's side. Even marketing can relate back to validation because it is easier for a well-established publisher to market something that they have validated. Each of these functions is interconnected in some way, but they have been categorized separately based on the theories they are drawn from. Moreover, formulating this framework as five

[163] M. Bhaskar, *The Content Machine*, p. 80.

separate functions is more practical and comprehensive for the following analysis. So, based upon the established framework, this thesis will analyse which functions of publishing preprints fulfil and what their role is within the system of academic publishing.

Chapter 3. Analysis

With both a comprehensive definition and elaboration of preprints and the theoretical framework established in the previous two chapters, this chapter aims to combine these two components in order to determine what functions of academic publishing preprints complement or have a significant role in. As established in the introduction, preprints are often referenced to as an important tool in academic publishing and they are even said to be 'becoming a common part of the scholarly publishing process'.[164] To evaluate the veracity of this claim and determine the specific functions to which preprints relevant role pertains, this chapter will undertake a detailed analysis of each function individually. By scrutinizing the role of preprints in relation to each function separately, this thesis can thoroughly assess the extent to which they fulfil a complementary role.

Rather than following the same structure as how the functions of academic publishing were presented in the previous chapter, this chapter will commence with the function in which preprints have the most clearly discernible complementary role, according to the following analysis. It may not come as a surprise that this function is dissemination, since this is considered one of the biggest selling points of preprints. Subsequently, the functions of editing and filtering and to what extent preprints have a complementary role in these functions will be elaborated upon. Lastly, this chapter will address the functions of validation and marketing and discuss why preprints have a smaller role in regard to these functions.

3.1 Preprints and the Function of Dissemination: A Complementary Role

The function of dissemination generally refers to making a work more widely available to a larger public, however in academic publishing this function is more complicated. As mentioned in the second chapter, dissemination is not merely about spreading a scholarly work to a wider audience but rather also about reaching the *right* audience. It is about disseminating a work to a disciplinary community, meaning a scholarly work has to reach a specific niche audience, which is much smaller.

[164] Puebla, Polka & Rieger, *Preprints: Their Evolving Role in Science Communication*, p. 64; Gross, 'Preprints and the Implications for Subsequent Peer Review and Publishing', p. 855; Penfold & Polka, 'Technical and Social Issues Influencing the Adoption of Preprints in the Life Sciences', p. 4; and Hoy, 'Rise of the Rxivs', 88.

However, whilst reaching this niche group of scholars is valuable, most scholars also find reaching a wider audience of equal importance. Academic publishers have a crucial role in reaching out to the niche group of scholars and ensuring they encounter the scholarly work, but this does not exclude the possible aim of reaching a wider audience as well. When authors publish their scholarly works, they may also desire a broad readership for their research. Moreover, not all academic papers fall into niche categories, and most are accessible to scholars within the general field, making it essential within the function of dissemination that scholarly works are spread as widely as possible to make it available to a large audience.

It is in this aspect that preprints fulfil a complementary role within the system of academic publishing. Preprints not only facilitate faster dissemination of scholarly works, but also provide an additional platform for readers to discover papers on specific topics. As a result, readers have an additional route through which to find the scholarly work, which has the potential to expand the audience beyond the one that would exist if the paper was solely published in an academic journal. This is already a common practice in field such as physics and mathematics, which have been using preprints for a longer time and search for scholarly works via preprint servers. By embracing preprints, scholars and publishers alike can profit from the benefits of wider dissemination.

This is logical in theory, and multiple scholars have noted that this does occur in practice as well. These scholars have observed that papers published in peer review journals that also have a preprint receive more citations and a higher Altmetric Attention Score than those without.[165] The Altmetric Attention Score is calculated with an automated algorithm which provides an indicator of the amount of direct attention a scholarly work has received online through mentions, discussions and shares, and where this attention is coming from.[166] While this score does not measure or indicate the quality of a work, it does provide an indication of how much attention a work draws. The Altmetric Attention Scores differ per scholarly discipline similarly to how the citation norms are different in each discipline.[167] The observation that papers that also have a preprint receive higher Altmetric Attention Scores indicates that preprints do indeed promote a wider dissemination.

---

[165] D.Y. Fu & Hughhey, J.J., 'Releasing a Preprint is Associated With More Attention and Citations for the Peer-Reviewed Article', *eLife,* 2019, pp. 1-12 (p. 5); and S. Serghiou & Ioannidis, J.P.A., 'Altmetric Scores, Citations, and Publication of Studies Posted as Preprints', *JAMA,* 319 (2018), pp. 402-403 (p. 402).
[166] 'How Is the Altmetric Attention Score Calculated?', *Altmetric,* < https://help.altmetric.com/support/solutions/folders/6000237502> (15 June 2023). For further reading on the Altmetric Attention Score see: *A Guide to Altmetric* by Cambridge UP's Author Hub.
[167] 'How Are Outputs Scored?' *Altmetric,* < https://help.altmetric.com/support/solutions/folders/6000237502> (15 June 2023).

Likewise, as mentioned, these papers that had preprints also received more citations.[168] Having more citations is another way to increase the dissemination of a scholarly work since readers will encounter it through another scholar's work. Kurtz et al., identified three factors that caused a citation advantage in this situation, namely the 'open access effect', the 'early access effect' and the 'self-selection effect'.[169] The open access effect means that there is an increase in the amount of readers as a result of wider accessibility; the early access effect refers to a work receiving more citations because it was available earlier, therefore longer, and has been read and cited more as a result; and the self-selection effect is a result of scholarly authors having a tendency to post work that is of the highest quality in order to acquire the most citations. Each of these effects can be linked to preprints and how they have aided in published, peer-reviewed papers receiving more citations when they have a preprint, which in turn leads to wider dissemination.

The EAVE II project provides us with an example from practice of how preprints could lead to the wider dissemination of the version of record. This project, which researched the severity of the Omicron variant of COVID-19 and the effectiveness of vaccines against it, was initially posted as a preprint, which resulted in it reaching the mainstream media headlines, including the BBC's.[170] These headlines were reported between the 22nd and 23rd of December 2021 and the preprint was posted in December 2021.[171] There is no exact date listed for when it was posted, however the article does mention that their study covered a time-period from November 1 until December 19, 2021, meaning that at the earliest this preprint was posted on 20 December 2021.[172] Not only media outlets but also specialised services such as the Science Media Centre, and national advisory groups, like the Australian Technical Advisory Group, took up on the research presented in the preprint and adopted it into their own research.[173] Within the first ten days, the preprint was downloaded 21005

[168] N. Fraser, et al., 'The Relationship Between bioRxiv Preprints, Citation and Altmetrics', *Quantitative Science Studies,* 1 (2020), pp. 618-638 (p. 619); Fu & Hughhey, 'Releasing a Preprint is Associated With More Attention and Citations for the Peer-Reviewed Article', p. 5; Serghiou & Ioannidis, 'Altmetric Scores, Citations, and Publication of Studies Posted as Preprints', p. 402.
[169] M.J. Kurtz, et al., 'The Effect of Use and Access on Citations', *Infortmation Processing & Management,* 41 (2005), pp. 1395-1402 (p. 1396).
[170] The article can be found in the BBC's Science Focus Magazine and is called 'How Worried Should We Be About the Omicron Variant?' by Dr. Jeremy Rossman (23 December 2021).
[171] 'The Power of Preprints: An Omicron Case Study', *The University of Edinburgh: Open Scholarship,* < https://libraryblogs.is.ed.ac.uk/openscholarship/2022/01/07/the-power-of-preprints-an-omicron-case-study/> (15 June 2023).
[172] A. Sheikh, et al., 'Severity of Omicron Variant of Concern and Effectiveness of Vaccine Boosters. Against Symptomatic Disease in Scotland (EAVE II): A National Cohort Study with Nested Test-Negative Design', *The Lancet: Infectious Diseases,* 22 (2022), pp. 959-966 (p. 959).
[173] 'The Power of Preprints: An Omicron Case Study', (15 June 2023).

times.[174] Due to the early access effect, the eventual article that was published in April 2022 was already cited many times and already had made headlines, meaning that the published paper had already had a head start compared to other papers in such journals that had not posted a preprint and had therefore not received any attention previously. Any statement on whether or not this paper would have received the same amount of attention if it had only been published in April 2022 is pure speculation. However, with how quickly the corona virus evolved, it would not be too far-fetched to consider that the Omicron variant of the COVID-19 virus, which the EAVE II project was centred around, was less of a hot topic four months later in April since a new subvariant had already taken its place.

It should be noted that this is an example in which time pressure played an important role, which is not the case in every discipline. HSS fields, for example, generally do not have the same time pressure on disseminating research, as was already touched upon in the first chapter. As was established earlier, the speed with which scholarly works are disseminated is not as relevant or valued as much in every discipline, which means that the extent to which preprints fulfil a complementary role in this aspect will differ per discipline as well.

Another aspect of dissemination as a function of academic publishing is the expansion of scholarly debate, which has been addressed in more depth in the second chapter. As discussed earlier, one of the key motivations behind the invention of preprints was to promote the discussion of research results before these were published.[175] The objective was to create a debate which would yield new insights and benefit the overall research while it was still a work in progress. Even when a preprint is already of the highest quality when it is posted, as per the self-selection effect that Kurtz et al. identified, the preprint still gives rise to scholarly debate preceding the formal publication, thereby expanding the debate beforehand, as exemplified with the EAVE II project.

In summary, preprints fulfil an important role by complementing the existing academic publishing system in regard to the function of dissemination. While preprints are not intended to replace the dissemination function of academic publishing, as they may not reach a specific niche audience as effectively, they do provide a valuable complementary tool for reaching a broader audience and expanding scholarly debates. By utilizing preprints in conjunction with formally published papers, the Altmetric Attention Score and citations of a scholarly work have shown a significant increase compared to papers without a preprint.

[174] 'The Power of Preprints: An Omicron Case Study', (15 June 2023).
[175] Teixeira da Silva, 'The Preprint Debate', p. 162.

3.2 Preprints and the Functions of Filtering and Editing

In relation to the functions filtering and editing, preprints can be argued to have a complementary role as well, albeit a less impactful one than on the function of dissemination. This subsection aims to analyse both these functions and examine the value of preprints in relation to them. Filtering will be discussed first, followed by an analysis of preprint's value to editing.

First, the function of filtering is about publishers making a curated selection of the works that are of a high enough quality and relevance to publish and preprint servers can play a part in this process. Aside from submissions, another way for publishers to obtain scholarly papers is, for instance, to go to conferences and invite scholars with appealing presentations to submit their future manuscript to a journal from their publisher.[176] This is relevant because it illustrates that publishers do not merely passively obtain scholarly manuscripts through submissions but that they also actively seek out innovative papers and attempt to acquire them for journals belonging to their publishing company. Preprint servers offer another platform for academic publishers to locate innovative papers that meet their standards of quality and relevance. Since preprints are openly accessible for everyone, publishers also have access to them and can curate the papers that they want to publish from these preprint servers.[177] While this does imply that the workload involved in the filtering process increases if publishers also have to look for potential journal submissions on preprint servers, it does appear that this is already happening in practice. As Penfold and Polka observe, there is a widespread practice of publishers inviting authors of preprints to submit to a journal.[178] However, most preprint servers do simplify the process of curation by illustrating metrics such as how many people have read a preprint and how often it was shared on social media. This could be an indicator for publishers about which preprints draw a readership and are popular.

Therefore, preprints can be used to complement the filtering function of academic publishing. By providing an additional platform where publishers can actively find innovative research, they can obtain new journal submissions that are current and meet the quality standards and scope of a journal. Additionally, publishers can gauge how a paper will be received due to the initial attention a preprint draws, which is reflected in its metrics. While

[176] Penfold & Polka, 'Technical and Social Issues', p. 4.
[177] Ibid., p. 4.
[178] Ibid., p. 4.

this is not an indispensable tool, it can prove to be a valuable addition to the academic publishing system.

Second, the editing function pertains to multiple editorial services that preprints do not fulfil, such as smaller editing to improve readability and writing style, peer review, and converting a scholarly work to the journal's house style. While preprints do not necessarily play a role in any of these services, however, there is an argument to be made for how preprints can still be useful to the editing function of academic publishing. The feedback that preprints receive, whilst undoubtedly considered useful in most instances, is not capable of replacing the rigorous peer review and editing services provided by publishers. Additionally, there is little literature available about the kind of feedback preprints receive, but it is often implied that this feedback is about ideas and insights regarding the research, rather than feedback on language use or grammatical errors.[179] These are services that the publisher provides and this is not the purview in which preprints can add value to the editing function.

Rather, there are two other manners in which preprints can add value to the editing function of academic publishing. The first draws on the aforementioned self-selection effect where authors, when their objective is for their preprint to do well, post a preprint of the highest quality. As a result of this effect, when authors intent for their preprint to also be formally published, their preprint is as close as it can be to the final version (of record). Nicholson et al. analyse the linguistic differences between the final preprint and the eventual paper published in a journal. They established that the linguistic difference between the preprint and published versions of a paper were minimal.[180] The main differences between these papers were due to converting the preprint into the publisher's house style before being published, and due to an 'increasing reliance on additional materials after peer review'.[181] This is likely due to the reviewers in the peer review process encouraging the use of certain additional sources to further strengthen the paper. Aside from these smaller differences, the preprint and published versions are not all that linguistically different, meaning that the preprint version was already of high quality before submission and needed less editing from the publisher.

The second aspect of preprints which is valuable to specifically the editing function of academic publishing is how preprints can lessen the time-pressure put on the publishers.

[179] Penfold & Polka, 'Technical and Social Issues', p. 4; Puebla, Polka, & Rieger, *Preprints,* p. 13 & 26.
[180] D.N. Nicholson, et al., 'Examining Linguistic Shifts During Publication' *PLoS Biology*, 20 (2022), pp. 1-22 (p. 16).
[181] Ibid., p. 11.

Generally, authors prefer for their work to be published as soon possible, which puts pressure on academic publishers to deliver.[182] As a result, many publishers have started making promises of shorter publication timelines.[183] However, one of the most time-consuming part of the publishing process is peer review, meaning that if publishers have to shorten their publication time, it is likely to come at the cost of time that is typically spend on peer review. Peer review, despite its contested nature, is necessary to ensure the quality of scholarly work, so rushing this process can mean a decrease in the quality of journal publications. Preprints can offer a solution to this issue since, according to Anderson, preprints function as a 'pressure valve for rapid publication outside journals'.[184] In other words, when authors release a preprint, it temporarily fulfils their desire to share their work publicly until publishers can complete the thorough peer review process and publish the paper in their journal. Therefore, this interim publication, to borrow Anderson's terminology, does add value to the system of scholarly communication as it plays a complementary role to the function of editing in academic publishing.

So, while preprints, in regard to both filtering and editing, do not necessarily appear to play as big a role within the academic publishing system as when it comes to dissemination, they do have additional value from which academic publishers can benefit. Preprints can therefore be considered a beneficial tool within the existing system.

3.3 Preprints and the Functions of Validation and Marketing

Lastly, in regard to the functions of validation and marketing, preprints seem to have the smallest role as part of the academic publishing system. The purpose of this subsection is to evaluate these two functions and assess to what extent preprints have a significant role in relation to them. This subsection begins by discussing validation and then delves into an analysis of the potential value of preprints in the function marketing.

First off, the function of validation in academic publishing entails that a scholarly work has gotten a publisher's stamp of approval because it was considered of a high enough

[182] Again, this is not the case in every discipline, but there are many scholars in multiple fields who share this view, which is why it has been included as an argument here.
[183] Penfold & Polka, 'Technical and Social Issues', p. 4.
[184] K. Anderson, 'The Tincture of Time – Should Journals Return to Slower Publishing Practices?', *Scholarly Kitchen*, < https://scholarlykitchen.sspnet.org/2017/03/28/the-tincture-of-time-should-journals-return-to-slower-publishing-practices/> (16 June 2023).

quality to be published after going through a critical filtering and peer review process, which does not align with the concept of preprints. Being formally published in a journal is a form of validation from the publisher, which assures the quality of a scholarly work. As per definition, preprints do not fit in this description and do not have a role in this function of academic publishing.

However, in the second chapter of this thesis, it was established that registration falls under this validation function as well, which is an aspect in which preprints can add value. As a result of a scholarly work being validated, it is also recognized when the author makes a discovery, meaning that due to the publication, an author can claim priority. It is with regard to the registration aspect of the validation function that preprints do have a useful role. As discussed in the first chapter, one of the advantages of preprints is that it allows for authors to establish priority. Every version of a preprint receives a timestamp, which can endorse that an author was the first to make a certain discovery. Therefore, preprints can have a relevant role when it comes to the validation function of academic publishing.

Nevertheless, the reason this is presented as a smaller complementary role in this chapter is due to the fact that establishing priority through preprints does not happen equally often across disciplines. As Vale and Hyman imply, not in every discipline are preprint established as a way to claim priority.[185] Where it is the norm in a field such as physics to claim priority through posting a preprint, there are other fields that still have to catch up, meaning that claiming priority through preprint still has to take hold.[186] In the case that preprints do allow an author to claim priority in their field, it does take away from the time pressure put on journals to publish a paper as soon as possible, which, as discussed in the previous subsection, is beneficiary to publishers. Thus, whether or not preprints have a complementary or relevant role regarding the validation function depends on what field a journal is published in and whether or not preprints are a recognized way to establish priority.

Lastly, the function of marketing does not initially appear to benefit from preprints, yet a closer examination reveals that there is an opportunity for academic publishers to employ preprints and preprint servers as marketing tools. The function of marketing entails that academic publishers aim to acquire, retain and expand their audience for a journal they publish in order to both generate revenue and increase their brand awareness. As established in the previous chapter, the objective of generating a revenue does not pertain to Open Access journals, yet increasing brand awareness is still relevant in these instances.

[185] R.D. Vale & Hyman, A.A., 'Priority of Discovery in the Life Sciences', *eLife*, 2016, pp. 1-5 (p. 5).
[186] Ibid., p. 5.

At first glance, preprints do not appear to have a significant complementary role in either of these marketing goals. Preprints are defined as being freely accessible scholarly manuscripts that have not (yet) been peer-reviewed. In this case, the emphasis is on the *freely* available aspect of preprints. Academic publishers cannot make a profit off of free preprints.

The only way in which preprints can financially benefit an academic publisher in this aspect is if a reader decides they want to read the peer-reviewed article based on the preprint. This then begs the question of whether or not academic publishers can use preprints as a marketing tool to convince readers to read the peer-reviewed article. Currently, academic publishers do not employ preprints in such way, but as publishers get more involved with preprints and preprint servers,[187] they might start using preprints as a marketing tool to generate a revenue.

Furthermore, in regard to increasing brand awareness, preprints do not initially appear to be a useful tool either, but upon closer inspection there is potential value in this regard as well. A preprint on its own is unable to increase brand awareness for a publisher, nor is it likely to draw in more authors to a certain academic publisher since the two are not linked to one another. Without a clear connection between a preprint and an academic publisher, preprints cannot function as marketing tools for a publisher who intends to promote brand awareness. However, as stated above, academic publishers are becoming more invested in preprints and their servers. Bigger publishers have bought or created their own preprint servers, which they can then employ to increase their brand awareness and thereby draw in more authors and readers. SSRN was bought by Elsevier in 2016 and is clearly linked to this publisher, therefore it can be used to promote brand awareness due to this connection.

Currently, preprint do not play a significant role in the marketing function, but there is a valuable opportunity for publishers to utilize preprints and preprint servers as marketing tools in the future as they get more invested in them. In short, where the role of preprints in regard to the marketing function is not all that significant now, there is a potential for it to become a useful tool in the future as academic publishers adopt preprints into their system.

In short, with regard to the functions of validation and marketing, preprints currently have the least relevance as a complementary tool. While the aspect of registration, which falls under validation, can be fulfilled by preprints, this depends on the discipline, meaning that there are cases in which preprints have no relevant role at all. As for marketing, preprints have the potential to be a valuable tool but in practice this potential is not fully realized at the

[187] Puebla, Polka, & Rieger, *Preprints*, p. 41.

moment. Nevertheless, I would argue that the role preprints have in marketing will grow in the near future as academic publishers grow even more involved in preprints and preprint servers and adopt them into their workflows.

In conclusion, preprints primarily serve a relevant or complementary role in the dissemination function among the five functions of academic publishing. Publishers find preprints to be an important and valuable tool in this regard. The functions of editing and filtering also benefit from preprints, albeit to a lesser extent compared to dissemination. However, preprints currently have a minimal role when it comes to validation and marketing functions. While preprints have the potential to complement the validation function through their registration aspect, the usefulness of this aspect depends on the discipline and its acceptance of preprints as a means to establish priority. Preprints also have potential to fulfil a relevant or complementary role in the marketing function but this has not been fully realized yet.

Conclusion

The objective of this thesis was to determine what role preprints have in academic publishing. As established in the introduction, many scholars have referred to preprints as becoming a common part of academic publishing and fulfilling a complementary or relevant role in this existing system. However, since there are multiple factors involved in this system, this thesis first had to identify the functions of academic publishing and built a framework with which to analyse to what extent these claims made by other scholars about the importance of preprints in academic publishing held up. In order to do so, five theories from other scholars about the functions of publishing in general or academic publishing specifically were compared in order to establish a framework that would allow an analysis of preprints in relation to academic publishing.

This new framework identifies five functions of academic publishing: filtering, editing, validation, dissemination, and marketing. When analysing the additional value that preprints have for each of these functions, this thesis determined that preprints hold the most relevance as a complementary tool to the function of dissemination. Then, both the functions of editing and filtering can also utilize preprints as a tool to optimize these functions. Even the validation function has the potential to benefit from preprints to some extent, depending on the discipline and its acceptance of preprints to claim priority. Concerning the marketing function, preprints possess the potential for playing a significant and complementary role. However, due to the current state of its adoption in academic publishing, preprints are unable to fully realize the extent of their potential impact at this moment. So, while it is clear that preprints have the most significant role with the function of dissemination, they do have smaller roles as well in other aspects of the academic publishing system.

Therefore, it can be concluded that preprints do, in fact, have a relevant role in academic publishing and that they function as a complementary tool that can add value to the scholarly communication process. This is further evidenced by publishers themselves who have gotten more actively involved with preprints servers, for instance by buying existing ones or starting their own, thereby proving that publishers also acknowledge the added value of preprints and their position within scholarly communication.

However, this thesis would like to emphasize that preprints can only fulfil a complementary role and are not capable of replacing formal academic publishing. As mentioned before, preprints are a valuable addition to academic publishing, but they cannot

fulfil every function of academic publishing. Even in terms of the dissemination function, preprints could not fully replace it, since it cannot reach the niche audience like academic publishers can. Preprints are only capable of complementing the existing system, not replacing it.

There is a reason why, in this entire thesis, preprints have not once been referred to as 'published', rather always in terms of 'posted' or some similar term. This is because there is an underlying question of whether or not preprints can be considered a form of publishing at all. In the current state preprints are in, I would argue that preprints cannot be considered publishing. Perhaps in the future as it continues to evolve, preprints can be considered a form of publishing on its own, but as it is now, preprints do not replace or subvert formal academic publishing. Rather, preprints complement the existing system and pose as a valuable tool, which publishers can utilize to enhance the services they offer to scholars.

That preprints can be a valuable tool for publishers has been made clear, but the same also goes for both authors and readers of scholarly works. For authors, posting a preprint, especially in its earlier stages, can result in receiving useful insights and feedbacks from readers, who in turn get to read about new developments in their field. In fields such as physics this is already a common practice, whereas most HSS fields have yet to adapt to preprints in such a way. Therefore, while I would certainly argue that preprints could be valuable to not only publishers, but to authors and readers as well, I do have to concede that there are significant disparities between disciplines. It is likely that it will take years before HSS adopts the use of preprints in the way STM disciplines have, but I am convinced that it will slowly spread to HSS fields too. The field of psychology has already begun using preprints due to its overlap with some STM fields, so it is a logical hypothesis to assume that preprint adoption will slowly spread to some extent to other social sciences that overlap with psychology. In this same manner, I do believe it will reach the humanities at some point as well, but not for another couple of years.

As publishers become increasingly involved with preprints and preprint servers,[188] it is likely that preprints are better integrated in the academic publishing system. The ways in which preprints fulfil a complementary role now will likely develop further if publishers fully adopt them into the services that they offer. However, while preprints are useful tools within the scholarly communication process, it is vital for scholars and publishers alike to remain critical of them. There are many advantages which preprints can bring to the scholarly

[188] Puebla, Polka, & Rieger, *Preprints,* p. 41.

communication system, but as the first chapter asserts, there are still many risks and concerns attached to the concept of preprints that should not be neglected. Ensuring that preprints have open data policies, meaning that all the data from the research should also be made available with the preprint, can help ensuring more transparency, but even then scholars have to remain critical. Preprints are a useful tool for discovering new developments within a field, but these papers should always be regarded with caution and not be cited without clarifying that it stems from a preprint.

Furthermore, do take into account that in this thesis, each identified function of academic publishing encompasses numerous smaller processes that have not been specifically addressed. Each function serves as an overarching term for the various aspects and roles that are associated with it. It is entirely plausible that preprints also play a complementary role in these smaller processes, or that they may bear no influence or relevance to them at all. This thesis has been focussed on the broader scope of academic publishing and has therefore not addressed these smaller processes. For example, the editing function encompasses many processes, such as peer review, converting a manuscript to the publisher's house style, and assigning a DOI to said scholarly work. For a future study, each individual smaller function could be analysed in more depth to determine the role of preprints there.

With this thesis, I have aimed to elaborate on the existing literature regarding preprints. The objective of this thesis was to build on what other scholars have stated, namely that preprints are important and complementary to academic publishing. Despite these statements proving to be true as a result of the analysis conducted in this thesis, none of these scholars had elaborated on the question of why exactly it was complementary and what functions it complements. By establishing a theoretical framework with which to analyse preprints in relation to academic publishing, I have been able to answer this question. Future scholars may build on this research using this new framework which fits within the digital age that preprints are a part of.

Bibliography

Anderson, K., ‘The Tincture of Time – Should Journals Return to Slower Publishing Practices?’, *Scholarly Kitchen*, < https://scholarlykitchen.sspnet.org/2017/03/28/the-tincture-of-time-should-journals-return-to-slower-publishing-practices/> (16 June 2023).

Anderson, R., ‘Revisiting – Journalism, Preprint Servers, and the Truth: Allocating Accountability’, *Scholarly Kitchen*, <https://scholarlykitchen.sspnet.org/2022/01/12/revisiting-journalism-preprint-servers-and-the-truth-allocating-accountability/> (13 June 2023).

Anderson, R. *Scholarly Communication: What Everyone Needs to Know* (Oxford: Oxford UP, 2018).

‘arXiv License Information’, *arXiv*, < https://info.arxiv.org/help/license/index.html>, (11 June 2023).

Bhaskar, M., *The Content Machine: Towards a Theory of Publishing from the Printing Press to the Digital Network* (Londen: Anthem, 2013).

Bhaskar, M., ‘Filtering, Framing, Amplifying: The Core of Publishing Now and Then’, *TXT*, 1 (2014), pp. 78-81.

Brierley, L., et al., ‘Tracking Changes Between Preprint Posting and Journal Publication During a Pandemic’, *PLoS Biology*, 20 (2022), pp. 1-22.

Clark, G. & A. Philips, *Inside Book Publishing* (Londen: Routledge, 2019).

Crick, T., et al., ‘A Multi-Disciplinary Perspective on Emergent and Future Innovations in Peer Review’, *F1000Research*, 6 (2017), pp. 1151-1215.

Fraser, N., et al., ‘The Relationship Between bioRxiv Preprints, Citation and Altmetrics’, *Quantitative Science Studies*, 1 (2020), pp. 618-638.

Fu, D.Y. & Hughhey, J.J., ‘Releasing a Preprint is Associated With More Attention and Citations for the Peer-Reviewed Article’, *eLife*, 2019, pp. 1-12.

Gross, J. & Ryan, J.C., ‘Landscapes of Research: Perceptions of Open Access (OA) Publishing in the Arts and Humanities’, *Publications* 3 (2015), pp. 65-88.

Gross, R.A., ‘Preprints and the Implications for Subsequent Peer Review and Publishing’, *Neurology*, 91 (2018), pp. 855-856.

‘How Are Outputs Scored?’ *Altmetric,* < https://help.altmetric.com/support/solutions/folders/6000237502> (15 June 2023).

‘How Is the Altmetric Attention Score Calculated?’, *Altmetric,* < https://help.altmetric.com/support/solutions/folders/6000237502> (15 June 2023).

Hoy, M.B., ‘Rise of the Rxivs: How Preprint Servers are Changing the Publishing Process’, *Medical Reference Services Quarterly*, 39 (2020), pp. 84-89.

Kaltenbrunner, W., et al., ‘Innovating Peer Review, Reconfiguring Scholarly Communication: An Analytical Overview of Ongoing Peer Review Innovation Activities’, *Journal of Documentation*, 78 (2022), pp. 429-449.

Klebel, T., et al., ‘Peer Review and Preprint Policies Are Unclear at Most Major Journals’, *PLoS ONE,* 15 (2020), pp. 1-19.

Kodvanj, I., et al., ‘Publishing of COVID-19 Preprints in Peer-Reviewed Journals, Preprinting Trends, Public Discussion and Quality Issues’, *Scientometrics*, 127 (2022), pp. 1339-1352.

Kurtz, M.J., et al., ‘The Effect of Use and Access on Citations’, *Infortmation Processing & Management*, 41 (2005), pp. 1395-1402.

McLuhan, M., *Understanding Media: The Extensions of Man* (New York: MIT Press, 1994).

Nicholson, D.N., et al., ‘Examining Linguistic Shifts During Publication’ *PLoS Biology*, 20 (2022), pp. 1-22.

Origgi, G. & G.B. Ramello, ‘Current Dynamics of Scholarly Publishing’, *Evaluation Review* 39 (2015), pp. 3-18.

Origgi, G., *Reputation: What It Is and Why It Matters*, (Princeton: Princeton UP, 2017).

Penfold, N.C. & J.K. Polka, ‘Technical and Social Issues Influencing the Adoption of Preprints in the Life Sciences’, *PLoS Genetics*, 16 (2020), pp. 1-16.

‘The Power of Preprints: An Omicron Case Study’, *The University of Edinburgh: Open Scholarship,* < https://libraryblogs.is.ed.ac.uk/openscholarship/2022/01/07/the-power-of-preprints-an-omicron-case-study/> (15 June 2023).

‘Preprints’, *Open Access Network,* < https://open-access.network/en/information/publishing/preprints#> (11 June 23).

‘Preprints in Europe PMC’, *Europe PMC,* < https://europepmc.org/Preprints> (10 June 2023).

Puebla, I., J. Polka, & O.Y. Rieger, *Preprints: Their Evolving Role in Science Communication* (California: ATG LLC (Media), 2021).

Roosendaal, H. E., & P.A.T.M. Geurts, *Forces and functions in scientific communication: an analysis of their interplay*, (1997), pp. 1-32. Paper presented at Conference on ‘Co-operative Research in Information Systems in Physics’.

Serghiou, S. & Ioannidis, J.P.A., ‘Altmetric Scores, Citations, and Publication of Studies Posted as Preprints’, *JAMA,* 319 (2018), pp. 402-403.

Servryugina, Y.V. & Dicks, A.J., ‘Publication Practices During the COVID-19 Pandemic: Expedited Publishing or Simply an Early Bird Effect?’, *Learned Publishing,* 35 (2021), pp. 563-573.

Sever, R., et al., ‘bioRxiv: The Preprint Server for Biology’, *bioRxiv,* 2019, pp. 1-19.

Sheikh, A., et al., ‘Severity of Omicron Variant of Concern and Effectiveness of Vaccine Boosters. Against Symptomatic Disease in Scotland (EAVE II): A National Cohort Study with Nested Test-Negative Design’, *The Lancet: Infectious Diseases,* 22 (2022), pp. 959-966.

Soderberg, C.K., Errignton, T.M., & Nosek, B.A., ‘Credibility of Preprints: An Interdisciplinary Survey of Researchers’, *Royal Society Open Science*, 7 (2020), pp. 1-17.

Strasser, C., ‘Preprints: The Bigger Picture’, *The Winnower*, 2016, pp. 1-6.

Texeira da Silva, J.A., et al., ‘Predatory and Exploitative Behaviour in Academic Publishing: An Assessment’, *The Journal of Academic Librarianship*, 45 (2019), pp. 1-8.

Teixeira da Silva, J.A., ‘The Preprint Debate: What Are the Issues?’, *Medical Journal Armed Forces India,* 74 (2018), pp. 162-164.

Turmudi, D., ‘English scholarly publishing activities in the industrial revolution 4.0: What, why, and how?’, *English Language Teaching Educational Journal,* 3 (2020), pp. 52-63.

Vale, R.D. & Hyman, A.A., ‘Priority of Discovery in the Life Sciences’, *eLife,* 2016, pp. 1-5.

‘What is Green Open Access?’, *Author Services: Supporting Taylor & Francis Authors,* < https://authorservices.taylorandfrancis.com/choose-open/publishing-open-access/oa-green-gold/> (10 June 2023).