

School of Psychology

Examining exaggerated claims in science communication

Luke Adam Bratton

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Supervisor: Petroc Sumner

**School of Psychology
Cardiff University**

November 2018

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed (candidate) Date.....

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD...(insert MCh, MD, MPhil, Ph.D. etc., as appropriate)

Signed (candidate) Date.....

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed (candidate) Date.....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organizations.

Signed (candidate) Date.....

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed (candidate) Date.....

Acknowledgments

Thank you to the ESRC and Cardiff University School of Psychology for providing funding to undertake this project.

I would like to thank my Supervisor Petroc Sumner for the great support through the years. And thank you to the rest of the InSciOut team for the amazing knowledge and advice.

Thanks to Aimeé for the snacks.

Thanks to Miller56 for the loud noises.

To Mum and Dad, thank you for always trusting me to do whatever I want to do.

To all of my friends from back in Cardiff, this type of undertaking is impossible without you.

Thanks Lins, for thanking me.

Thanks Jaakko, for the inspirational phone calls.

Carolina, thank you for being here, even when you were 5000 miles away.

Abstract

This thesis is concerned with investigating the exaggeration of health-related research in the media. Typically, research findings published in peer reviewed journals are transmitted to the news via press releases created by universities and journal press offices. Research has shown that exaggeration of key aspects of the research relevant to the health-related behaviour of readers is often exaggerated in the news. Observational research has shown that the presence of exaggeration in press releases is related to exaggeration in the news (Sumner *et al.*, 2014).

Firstly, I report my largely successful replication of this key study using more recent retrospective observational data. I show that discourse on openness in animal research and exaggeration of findings is linked to positive changes in science reporting.

The study in chapter three compares data collected before versus after the release of Sumner *et al.* (2014) to detect any change the reporting of research findings following the release of this high profile paper. Between the sample periods, exaggeration in press releases had reduced, suggesting that press officers had become more cautious in their reporting of research findings.

Chapter four describes a randomised controlled trial which directly modified the output of press offices to observe whether press release content had a direct effect on news. A high level of condition non-adherence meant that this “per protocol” comparison was not possible. An “as treated” analysis demonstrated the same relationship between press releases and news articles as in the replication in chapter two, and the comparison in chapter three.

Chapter five reports a study which aimed to test the influence of press release content on news selection and content using experimental methods. Journalism students were given identical press releases which were modified between participants and asked to select those which were newsworthy, and to write a news article based on one of the press releases. Article topic significantly predicted the proportion of stories selected as newsworthy, whereas manipulations to press release content did not. There was no difference in the content of participants' news articles. Since this did not replicate the observational results of previous chapters, the experimental setup may have not been an accurate homologue of the varied real-world journalistic environment, as students were probably behaving pedagogically, and were acting in response to the same instruction.

To see whether the instruction given to participants could influence their output, undergraduate psychology students were asked to rewrite articles in a way that is either more concise, more appealing, or accurate to the source. Variation in participants' output was not explained by instruction, rather it was again explained by the topic of the research. As participants' free text comments from chapter five suggest, this may be because the perceived level of interest in the story is most important.

Finally, I draw conclusions relating to the improvement of reporting in the science media process. It appears that there is no penalty for accurate reporting, and the inclusion of important scientific details in health-related press releases. Press officers should therefore follow the guidelines of the Academy of Medical Sciences and the Science Media Centre.

Contents

CHAPTER 1 – GENERAL INTRODUCTION

1.1. Dissemination of science	1
1.1.1. Introduction	1
1.1.2. Theories of science communication	
1.1.2.1. <i>Methods of outreach</i>	2
1.1.2.2. <i>Justification for outreach</i>	3
1.1.2.3. <i>Justification for outreach</i>	5
1.1.2.4. <i>The public need for information</i>	6
1.2. People search for health information online	8
1.3. The scale of health-news	8
1.4. Change of behaviour	10
1.5. Science-news controversy	11
1.5.1. <i>A note on ‘exaggeration’</i>	12
1.6. Inaccuracies in news	15
1.7. The Journalistic Environment	16
1.8. The role of the press release – key studies	20
1.8.1. <i>Schwartz, Woloshin, Andrews, and Stukel (2012)</i>	20
1.8.2. <i>Sumner et al., (2014)</i>	21
1.8.2.1. <i>Justification</i>	21
1.8.2.2. <i>Exaggerated statements of relationship</i>	22
1.8.2.3. <i>Exaggerated inference from non-human research</i>	23
1.8.2.4. <i>Exaggerated advice</i>	24
1.8.2.5. <i>Findings</i>	24
1.8.2.6. <i>Interpretation</i>	24
1.9. Synopsis	25
CHAPTER 2 - REPLICATION	
2.1. A note on contribution	28
2.2. Introduction	28
2.3. Method	
2.3.1. <i>Data collection</i>	35
2.3.2. <i>Article Coding</i>	37
2.3.2.1. <i>Coding of advice</i>	39
2.3.2.2. <i>Coding of statements of relationship</i>	39
2.3.2.3. <i>Coding of samples</i>	40
2.3.3. <i>Analysis</i>	41
2.3.3.1. <i>Advice</i>	42
2.3.3.2. <i>Statements of relationship</i>	43
2.3.3.3. <i>Inference from non-human samples</i>	43
2.4. Results	
2.4.1. <i>Press release exaggeration</i>	44

2.4.2. <i>Exaggeration in news relative to exaggeration in press releases</i>	45
2.4.3. <i>News uptake relative to exaggeration in press releases</i>	48
2.5. Discussion	50
2.5.1. <i>Differences to Sumner et al. (2014)</i>	51
2.5.2. <i>Limitations</i>	53
2.5.3. <i>Conclusion</i>	57
CHAPTER 3 – INTERRUPTED TIME SERIES	
3.1. Introduction	60
3.1.1. <i>Paper impact</i>	60
3.1.2. <i>Causal language</i>	61
3.1.3. <i>Categorising statements of relationship</i>	63
3.1.4. <i>Present study</i>	66
3.2. Method	67
3.2.1. <i>Data collection</i>	67
3.2.2. <i>Article coding</i>	69
3.2.3. <i>Analysis</i>	70
3.3. Results	
3.3.1. <i>Journal article exaggeration</i>	72
3.3.2. <i>Press release exaggeration</i>	73
3.3.3. <i>News exaggeration</i>	73
3.3.4. <i>News exaggeration as a function of press release exaggeration</i>	74
3.4. Discussion	75
3.4.2. <i>Limitations</i>	78
3.4.3. <i>Conclusion</i>	79
CHAPTER 4 – RANDOMISED CONTROLLED TRIAL	
4.1. A note on contribution	82
4.2. Introduction	82
4.3. Method	
4.3.1. <i>Participants</i>	84
4.3.2. <i>Design and procedure</i>	85
4.3.3. <i>Article coding</i>	86
4.3.4. <i>Analysis</i>	87
4.4. Results	
4.4.1. <i>Intention-to-treat analysis – statements of relationship</i>	88
4.4.2. <i>Intention-to-treat analysis – caveats</i>	89
4.4.3. <i>As-treated – statements of relationship</i>	89
4.4.4. <i>As-treated – caveats</i>	90

4.4.5. <i>Condition non-adherence</i>	91
4.5. Discussion	91
CHAPTER 5 – EXPERIMENTS WITH JOURNALISTS	
5.1. Introduction	94
5.2. Method	
5.2.1. <i>Participants</i>	96
5.2.2. <i>Materials</i>	97
5.2.3. <i>Design</i>	
5.2.3.1. <i>News selection</i>	99
5.2.3.2. <i>News content</i>	101
5.2.4. <i>Procedure</i>	102
5.3. Results	
5.3.1. <i>News selection</i>	104
5.3.2. <i>Analysis of comments – reasons for news selection decisions</i>	106
5.3.3. <i>News content</i>	109
5.3.4. <i>Other sources used</i>	110
5.4. Discussion	112
5.4.1. <i>News selection</i>	112
5.4.2. <i>News content</i>	115
CHAPTER 6 – INSTRUCTION EXPERIMENT	
6.1. Introduction	119
6.2. Method	
6.2.1. <i>Participants</i>	121
6.2.2. <i>Materials</i>	121
6.2.3. <i>Procedure</i>	122
6.2.4. <i>Design</i>	123
6.2.5. <i>Analysis</i>	123
6.3. Results	124
6.4. Discussion	126
CHAPTER 7 – GENERAL DISCUSSION	
7.1. Overview	131
7.2. Interpretation	133
7.3. Generalisation	138
7.4. Observational research	139
7.5. Concordat on accuracy	142
7.6. Conclusion	144
References	146
Appendices	154

CHAPTER ONE – GENERAL INTRODUCTION

1.1. Dissemination of science

1.1.1. Introduction

Any scientist who publishes research today has to accept that the majority of public exposure to their research findings will be mediated via the keyboard of a journalist. For the general public, research findings are mostly inaccessible to anyone who cannot afford to pay the hefty subscription fees to journals. Even if they do manage to get hold of an open access research paper, the contents will probably be too technical and impenetrable to understand. An Ipsos MORI survey of 1749 UK adults aged 16-years and over found that 55% of people agree that science is too specialised for most people to understand (Castell *et al.*, 2014). In addition, research findings printed in journal articles are written with an inductive writing approach, where the key conclusions are buried in the article after significant qualification. This writing style introduces key concepts and provides important definitions, before developing narrative and describing findings that justify a conclusion. By contrast, news is presented in a far more easily accessible ‘inverted pyramid’ writing style where the key conclusions are presented at the start of the text, or in the title (Pottker, 2003). In this writing style, the text starts with a lead sentence that summarises the key message. This lead sentence should answer as many of the “W questions” as possible (when, where, what, who, and why), to provide as much of the information of interest to the reader. The information following the lead sentence is less important, and is presented in order of decreasing relevance. Compared to the scientific writing style, the inverted pyramid news writing style raises a few potential issues when presenting scientific research. Firstly, the inverted pyramid style may have

become popular because the headline and the lead sentence have greater communicative potential - to grab attention even for stories that are uninteresting to the reader (Pottker, 2003). This limits the ability to adequately describe complicated scientific findings that might require more than just one sentence. The shortest form of scientific communication - the abstract – handles this task by providing multiple sentences of background and qualification before providing a sentence of conclusion comparable to a news lead sentence. But even scientific abstracts have been found to contain unjustified conclusions (Yavchitz *et al.* 2012). Secondly, the heavy focus on the lead sentence in the inverted pyramid writing style means that details that are important to understand the claim being made, such as caveats regarding the study design, are provided further down the article body, and can be missed by readers. This means that readers may be presented with information without justification; without the extra information needed to understand how the information in the headline or lead sentence was deduced.

1.1.2. Theories of science communication

1.1.2.1. Methods of outreach

The majority of scientific findings that are disseminated to the public are done so through the news media. This type of relationship between science and the public is a form of scientific outreach, where information passed from scientists is essentially translated from the technical nature of science writing into articles that are more understandable to the layperson. But outreach can also take the form of events or organised by scientists, where the public are presented with workshops or activities that aim to engage and inform. Such events can engage

parties in two-way communication in a way that newspapers and television news cannot. However, the negative of face-to-face science communication is that it is more time consuming and costly for a limited audience of tens to hundreds, as opposed to the thousands or even millions that can be reached by traditional news media (Bultitude, 2010). Social media is a more recent addition to the outreach toolbox. Scientists such as Brian Cox can instantly broadcast information to millions of people with very little effort. At the time of writing, Brian Cox has over 2 million followers on Twitter, which is currently half a million more than The Daily Mail Twitter account. Anyone can create a social media account to disseminate research, but very few people can gain such an outreach potential. Also, given the lack of credibility of social media, it is difficult for a scientist to gain the trust of online viewers. This is why the traditional channel of dissemination of information through news media is still so successful.

1.1.2.2. Justification for outreach

When a new piece of information is discovered through the scientific method, scientists engage in the prescribed form of communicating findings via peer-reviewed publication in journals. Whilst this process is becoming increasingly accessible to the public view, published scientific research articles are still complicated and difficult to understand for the public. But why do scientists need to perform outreach activities with their research at all? Research Councils UK (2010) recommends that researchers engage in outreach for a number of personal career benefits, including skills development, improving one's personal profile, and potential to form new collaborations and gain funding. Research Councils

UK, since renamed UK Research and Innovation, require that researchers demonstrate impact and outreach activities as a condition of their funding. Much of the funding for UK science activities comes from the public funds, so regardless of whether there is personal benefit, or institutional requirement, public outreach could also be seen as a duty to the public.

Outside of the personal benefits and institutional requirements, there has been a great deal of discourse surrounding the need for outreach via science communication for the public good. In the 1980's, the emerging consensus from social scientists was that the public were sceptical about science because their lack of scientific knowledge (Dickson, 2005). As a solution to this, the suggestion was that providing a higher volume of high quality science news could improve the public lack of knowledge and overcome their scepticism. In this 'deficit model' of science communication, the main objective became to produce more, high quality science news content, and make it available to the public (Dickson, 2005). By the early 2000s, the deficit model had become largely discredited because accurate scientific information was not shown to increase public trust in science (Sturgis & Allum, 2004). Take for example the case of the NASA scientist David Morrison, who claims that his research activity has been disrupted a number of times due to the need to answer questions from the public regarding the theory that a fictional planet has been prophesied to collide with Earth. Despite repeatedly providing scientific evidence that such a planet does not exist, the phenomenon remains (Selk, 2017).

It appears the public do not base their decisions simply on accurate scientific information; rather they base decisions also on religious, cultural, ethical, historical, and personal concerns (Brown, 2009). In other words, a consideration of the individual differences between people has changed the target of science communication from being directed at ‘the public’ as a logical entity that simply needs to digest more information, to being aimed at a diverse and dynamic public (Einsiedel, 2007). Therefore the focus of science communication has moved away from pure dissemination, towards dialogue (Felt, 2003; van der Sanden, & Meijman, 2008), to accommodate the vast differences in public perception. This further justifies other areas of outreach, such as social media, multimedia and technology, and university-organised outreach events and public conferences.

1.1.2.3. Justification for accuracy

Given that the consensus is that the deficit model is no longer relevant, and that a dialogue model (public involvement in discourse), and a participation model (public participation in science-related activities) of science communication has been emergent, consideration should still be given to the need for accurate science communication. Clearly, providing scientific information to the public cannot alone improve understanding of science, in the same way that a traditional chalk-on-blackboard math lesson is not suitable for all learners. This should not mean that the accuracy of science communication is relaxed in favour of strategies to make scientific information more engaging. Whilst science communicators need to create engaging content that can be shared in engaging ways, they also need to ensure they act ethically in their reporting. The Society of Professional Journalists published a journalistic code of ethics that highlighted

the need for accuracy in reporting, even at the cost of time and article format (Society of Professional Journalists, 2014). This is because accurate information should still be an underlying constant in the light of engaging methods of communication, arguably even more so, since an increase in consumption of engaging content could otherwise be a driver in the exposure to inaccurate information. As discussed in later sections in this introduction, inaccurate information can have devastating consequences.

There is evidence that the ethical concerns of journalists do occur in practice. In the severe acute respiratory syndrome (SARS) outbreak in China, the role of journalists appeared to change from that of reporter to that of public servant (Wilkins, 2006). Reporting heavily focused on the disease and its mitigation, and became more factual and informative, before following the pattern of returning to normal following the decline in deaths from the disease (Aldeman, & Verbuge, 2000). Though this is an extreme example, it does show that there are underlying ethical considerations in science communication that can manifest as increased accuracy.

1.2.2.4. The public need for information

In science communication, as with news in general, it is best practice for stories to be published whilst they are still current. This can cause problems in science communication because scientific findings are rarely breakthroughs; rather, findings build upon previous work to add information and expand knowledge in the field. This process means that evidence discovered today may be superseded by new evidence tomorrow, which means that science news presented to the

public can appear contradictory or simply incomplete. In a story about disintegration of ice sheets, uncertainty about the data meant that it was published in a way that could misinform readers allowing them to underestimate the risk of sea-level rise (Keohane, Lane, & Oppenheimer, 2014). Whilst it could be stated that science communicators have a duty to disseminate such information to the public to allow them to use it to make informed decisions, for example about their own emissions, it could be argued that it is unethical to provide information that could lead to negative outcomes. The Society of Professional Journalists (2014) code of ethics states that journalists should balance the public's need for information against their potential harm, and to recognise that simply having access to information does not provide an ethical justification for publication. This also raises the question of whether there are differences in scientific fields, or other characteristics about science stories that make them more or less suitable for the public (Medvecky, & Leach, 2017). For example consider the following two findings: 'the consumption of almonds can modulate mood', and 'a certain chemical has been found to have the appropriate thermal efficiency for use in the propulsion of warheads'. It is debatable that if the following findings were published on the same day, they are unlikely to be seen as equally suitable for publication as news stories. Both sound like they could be interesting news stories, but there is a difference between the two articles in the balance between the public need for information, and the ethical considerations of publishing such information. News related to health-related research findings represents a case where the public need for information and the ethical justification are strong. As is demonstrated below, the public also frequently seeks health news.

1.2. People search for health information online

In the USA, the most common sources of information for new science research findings were television news programmes (42% of respondents), other television programmes (26%), and print newspapers (23%) (Castell *et al.*, 2014). Taking just the data for the youngest 510 respondents (aged 16 to 24), the second and third most frequently used sources were online newspapers or news websites (24% of respondents), and social networks (21%). So generally, people are exposed to research findings through more traditional sources, but the shift to online sources in younger people suggests that the Internet will be more frequently used for obtaining scientific findings in future. These results also indicate that news media outlets have a large potential for influence, since it is their content being presented through most of the top information sources.

1.3. The scale of health-news

More people go online for health information everyday than go to see health professionals in the US (Fox & Rainie, 2002). For the year 2000, it was estimated that 52 million American adults relied on the Internet to make their health decisions – for 2002 this estimation had risen to 73 million, and in 2006 it was an estimated 80% of US Internet users, or 113 million adults (Fox, 2006). To put this into perspective, the first iPhone was launched one year after the data collected for this estimate, so it is now likely to be a much higher number.

Smartphone users more frequently access the Internet than computer users (Zach, Dalrymple, Rogers, & Williver Farr, 2012), and accessing health information had become the third most popular use of the Internet in all those aged 18 years and older (Zickuhr, 2010). In recent years, it has been noted that excessive use of the

Internet to search for health related information has been associated with anxiety, and this has been coined “cyberchondria” (Starcevic & Berle, 2013).

Health was the 8th most commonly reported news topic in the US between January 2007 and June 2008 accounting for 3.6% of all coverage, ahead of business, lifestyle, and sports news – just 5 months before the US Presidential election, which accounted for over 20% of news stories. Specific diseases such as heart disease and cancer are the most frequent health news topics, at 41.7% of health news coverage. Cancer accounted for 10.1% of all coverage. Evening television news (8.3%) and newspapers (5.9%) were the most frequent mediums to report on health news. Online news by comparison only reported on health issues 2.2% of the time, but the overall volume of online news is much higher (The Kaiser Family Foundation, & The Pew Research Centre’s Project for Excellence in Journalism, 2008).

Given the vast scale of the public’s potential reliance on health news for information, it is worrying that around 75% of online health seekers responded “only sometimes”, or “hardly ever/never” to a question about whether they check the source of health information they find online. The Society of Professional Journalists (2014) state in their code of ethics that journalists should take responsibility for the accuracy of their work, but as is discussed below, health news is not always accurate.

1.4. Change of behaviour

In addition to finding so much of their health news online, people also change their behaviour based on what is reported. The advertising industry is built on trying to influence the behaviour of the consumer, and health-related behaviour has been seen to change in conjunction with targeted media campaigns. In 2000, anti-smoking TV campaigns aired in some areas across the United Kingdom were associated with higher rates of smokers quitting and lower rates of ex-smokers relapsing, in comparison with areas where the campaigns were not aired (McVey & Stapleton, 2000). A review of health-oriented mass media campaigns found that active campaigns were associated with concurrent or subsequent changes in behaviours related to the focus of the campaign (Wakefield, Lokin & Hornick, 2010). For example, a news campaign surrounding the World Transplant Games Federation was associated with an increase in organ donations in cities where the campaign was active, but these increases were not sustained when media exposure stopped (Slapak, 2004). Behaviour change has also been measured in relation to news reporting directly on published research. In Australia, sales of iodised salts increased in the weeks immediately following news reports of a study highlighting the issues with iodine deficiency, and the benefits of consuming iodised salt (Li, Chapman, Agho, & Eastman, 2008). Similarly, in a review of 20 interrupted time-series studies, five examining news reports of health findings and a further 15 reporting on mass media campaigns, all were found to be related to a change in their related outcome measures (Grilli, Ramsay, & Minozzi, 2002). The studies examining news coverage of health related-issues found that following coverage there was a reduction in incidence of Reye's Syndrome (Soumerai, Ross-Degnan, & Kahn, 1992), an increase in use

of HIV counselling services (Tesoriero & Sorin, 1992), an increase in use of mastectomy in breast cancer patients (Nattinger, Hoffmann, Howell-Pelz, & Goodwin, 1998), a reduction in the use of calcium channel blockers (Maclure *et al.*, 1998), and a reduction in hysterectomy rates (Domenighetti *et al.*, 1988).

1.5. Science-news controversy

The studies reviewed by Wakefield, Lokin, and Hornick (2010), and Grilli, Ramsay, and Minozzi, (2002) generally reported cases where campaigns and interventions were associated with positive changes, or a reduction of negative changes in health related behaviours. But perhaps the most high profile cases of media impact are the controversies related to negative changes. In the most famous example for health-related findings, the discredited link between vaccines and autism remains a persistent perspective in the media (Poland & Spier, 2010). There have been cases of unvaccinated individuals causing outbreaks of measles in populations where measles had been previously eliminated. In one case, an unvaccinated individual triggered an outbreak of a strain of the virus that spread to 34 others. Of these infected individuals, 31 had previously declined vaccination primarily due to concerns of the vaccines adverse effects (Parker *et al.*, 2006). In another case in 2008, a single unvaccinated child infected 11 others with measles; the parents of the majority of the unvaccinated children cited concerns with adverse effects (Sugerman *et al.* 2010).

The origin of this rekindled concern with adverse effects of vaccines was the press relations follow-up to a now retracted study that tentatively suggested a

link between vaccinations and a syndrome characterised by a bowel disorder and cognitive issues. The study used a small sample of self-selected participants, with a confounded self-report outcome measure, and was later found to contain fraudulent data, and was led by an author with a major conflict of interest (Godlee, Smith, & Marcovich, 2011). The original report also actually made the conclusion that it “did not prove an association between measles, mumps, and rubella vaccine and the syndrome described. [...] Published evidence is inadequate to show whether there is a change in incidence or a link with measles, mumps, and rubella vaccine”. Despite this, an Internet search for the exact phrase ‘vaccines cause autism’ will no doubt yield recent articles reporting this fabricated statement of relationship. The subsequent press release from The Royal Free Hospital School of Medicine (Hutchinson, 1998) did not actually exaggerate these claims, but the subsequent press tour of the principal author of the original study contained discussions of a “gastrointestinal origin of autism” linked to “damage caused by the MMR vaccine” (Autism Network for Dietary Intervention, 1998).

1.5.1. A note on ‘exaggeration’

The term ‘exaggeration’ will be used throughout this thesis, and although it is discussed later in this chapter, and in further chapters, this is a good example to expand upon. Exaggeration in science media can be defined in multiple ways depending on the context – such as cases in news in comparison to what the press releases said, or in comparison what journal article did. It could be said that a news article reporting on a study of mice could be exaggerated if it stated a recommendation for humans, or simply if the news article was published with a

photo of a human whilst reporting on findings in mice. Perhaps the most obvious form of exaggeration is that of statements of relationship (as in the above case of the vaccine scare). That is, when the relationship between two variables is overstated in comparison with the source article. For example, if a journal article states that ‘a sedentary lifestyle is related to increased apathy’, a news article could be deemed to have exaggerated if it stated that ‘sedentary lifestyle leads to apathy’. The correlational language - ‘related’ - is exaggerated to be causal - ‘leads to’. If an observational study reported that ‘vitamin D reduces fatigue’, this would be an exaggeration of the inference that is justified by the study design. If a news article then went on to say ‘fatigue is cured by vitamin pills’, this could be classed as exaggeration of what the study did, but not what the study said.

This interpretation uses the term ‘exaggeration’ to label instances where the change in information from one article to another is unjustified. It is noted from personal conversations with press officers that ‘exaggeration’ could be seen as a negative term, potentially implying that an inflated claim had been written on purpose. In this thesis, exaggeration merely defines information in excess of its source.

In previous research there have been a number of different ways of interpreting the differences between articles in science communication. Schwartz, Woloshin, Andrews, and Stukel (2012) based their interpretation on the presence or absence of quality measures in health-related news reports, in comparison to their source press releases. The quality measures used were the presence of basic study facts,

study limitations, the main results, and harms of interventions. A news article containing such details would be rated as higher quality than an article without such details. For comparisons of the results of the research, this method only allows for recording the presence or absence of a quantified result, and whether the result was quantified with the correct statistic, but it does not allow for a comparison of different levels of the same information, such as the different levels of a relationship between variables described previously. This limitation is only due to the quality measure of Schwartz, Woloshin, Andrews, and Stukel (2012) being a binary measure. The term ‘quality’ can actually be applied to exaggeration, given that a lack of exaggeration in a news article would classify the article to be high quality. ‘Exaggeration’ is simply a binary interpretation of the comparison between articles.

Similarly to Schwartz, Woloshin, Andrews, and Stukel (2012), Schwitzer (2014) assessed news articles on whether they adequately covered each of ten criteria such as quantifying benefits, evaluating quality of evidence, and whether the article quoted independent sources. This is similar to the coding of Schwartz, Woloshin, Andrews, and Stukel (2012), and could be interpreted as measures of quality. Schwitzer (2014) focused purely on news about new medical tests and equipment whereas this thesis is concerned with all news reporting research findings relevant to human health, with a main focus on the accuracy of the main statement of relationship, the sample stated, and advice given. In this regard, the methods of Schwitzer (2014) and Schwartz, Woloshin, Andrews, and Stukel (2012) are both too restrictive, as some of the assessments would not be relevant.

1.6. Inaccuracies in news

It is unlikely that health-news readers are frequently exposed to scandals of the magnitude of the vaccine controversy, but unintentionally inaccurate stories, or stories lacking important scientific details are common. An analysis of 2050 health news stories reporting on health-interventions (such as reports of new drugs) scrutinised each article for 10 features that were considered to be important for readers to make informed decisions (Schwitzer, 2015). The features scrutinised were whether the article covered costs, benefits, and harms of intervention, and whether it evaluated the quality of evidence, widened the diagnostic boundaries of the treatment, quoted independent sources, compared the treatment to others, mentioned availability of the intervention, assessed treatment novelty, and whether it relied solely on the press release. These features were selected as those that are seen as the most important pieces of information when reporting on new treatments. Over 60% of news articles did not satisfactorily quantify harms, benefits, and costs of the intervention. An independently developed news quality rating system was developed to analyse health advice in Australian magazines and newspapers (Wilson, Bonevski, Jones, & Henry, 2009). This rating system was similar to that used by Schwitzer (2014) and Schwartz, Woloshin, Andrews, and Stukel (2012) in that it listed specific features that, if missing, would indicate a low quality news report. Some of the criteria were identical to those used by Schwartz, Woloshin, Andrews, and Stukel (2012), and Schwitzer (2014), such as mentioning harms of treatment, and whether the report was based on anecdotal evidence, but with additions such as the article containing advertisements, and the author having a conflict of interest. Using this rating system to create a satisfaction score for each article, it was

shown that the highest average score was 58%, attained by broadsheet newspapers, but all other news sources scored less than 50% satisfaction (Wilson, Bonevski, Jones, & Henry, 2009). So what is the reason for such seemingly low quality news? To try to answer this question we need to consider the environment in which journalists operate.

1.7. The Journalistic Environment

In an investigation of the number of newsroom employees, newsroom revenues, and the quantity of newspaper content produced between 1985 and 2004, Lewis, Williams, & Franklin, (2008a) showed that whilst resources and staff had remained fairly constant, total newspaper content more than doubled over the same period. This finding is purely based on physical newspaper content; the inclusion of growth in online news would likely make this figure much higher. In a linked survey of 42 journalists, the majority stated that they felt that they had to produce more content than they had to a decade previously. The majority of these same journalists also indicated that they use press relations material sometimes, or often, to inform their stories, with health-news stories being the most common to be informed by press releases (Lewis, Williams, & Franklin, 2008a).

Press releases are summaries of more complex events or findings produced by public relations employees and are integral in the dissemination of research findings (figure 1.1). A press release is often written in conjunction or under consultation of the original author of the research paper, but typically using the similar inverted pyramid style used by journalists, which is a more easily

digestible format than the inductive style used by journal articles (Pottker, 2003). Scientific journals and universities frequently produce press releases to inform the media about new research findings. In an analysis of 90 news reports based on research published in the *Lancet* and *The BMJ* medical journals, around 80% had been initially issued as a press release by the journal (Entwistle, 1995). Press releases are now so commonly used that journalists report that the daily task of sifting through their email inboxes for stories amongst press releases is a time consuming task in itself (Williams & Clifford, 2009). A study of 53 local television health-news reporters in United States showed that their news stories are most frequently motivated by the direct contact from a public relations spokesperson (~50%), or by a press release (~45%). By contrast, only around 20% of the respondents indicated the medical journal as a source of motivation (Tanner, 2004).

The compounding of more work for journalists to do, in less time, with fewer resources, makes the press release an attractive resource. Davies (2009) suggested that this environment had led to what he coined “churnalism”. Churnalism is a neologism combining ‘journalism’ and ‘churn’ to suggest the practice of journalists churning out articles at high frequency. The busy newsroom environment, the requirement to create new content as frequently as possible, the short turnaround times for new articles, and the availability of press releases as information subsidies which are seen to contain most of the information required for a news report, means that journalists have been accused of a greater focus on information reproduction rather than curation (Davies, 2009).

Regarding health-related research, churnalism is the heavy reliance on university and publisher press releases, to create newspaper articles. Indeed, the synthesis of the term 'churnalism' was based on the finding that the content of around 40% of news stories reporting on health and nature are entirely or mainly reliant on the content of press releases (Lewis, Williams, & Franklin, 2008b). If churnalism is a persistent practice, it would follow that there would be a high proportion of similarity between press release content and news content. This supports the findings of Schwartz, Woloshin, Andrews, and Stukel (2012), that there is an association between the incidence of their quality measures in press releases and news. Although, the news articles did not always state the features mentioned in the press release, and news articles were found to sometimes contain relevant information that was not present in the press release. This suggests that churnalism, in practice, is not a simple copy and paste of information from press releases. Rather, journalists will be relying on press releases for a great deal of information, but they will supplement this with information sought from other sources, or from their own interpretation.

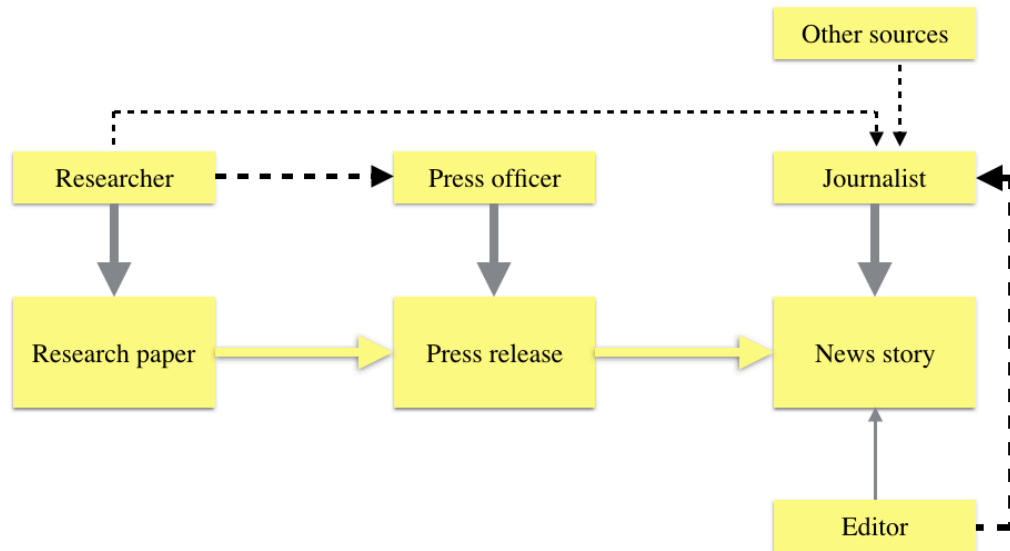


Figure 1.1. The typical science news process. Research papers inform press releases, which in turn inform news stories. Grey arrows indicate authorship. Black dotted arrows indicate the dissemination of information in person-to-person interactions.

However, if the quality and accuracy of news is generally thought to be low, but there is at least some level of reliance on press releases, it follows that focus should be directed towards the quality and accuracy of press releases. In data reported above, Schwitzer (2015) showed that a large proportion of news articles reporting on health interventions miss out important details. But in the same analysis, it was shown that inclusion of these important features was as unsatisfactory, if not worse, in press releases reporting on the same stories. The content of press releases, and their role in the communication of accurate information is the focus of the studies by Sumner *et al.* (2014), Schwartz, Woloshin, Andrews, and Stukel (2012) described below, and the main focus of this thesis.

1.8. The role of the press release – key studies

The interaction between the press release and the news article is clearly important in understanding the presence of inaccuracies in health reporting.

The following are summaries of key empirical studies investigating the relationship between information contained in press releases, and the same information contained in the news.

1.8.1. Schwartz, Woloshin, Andrews, and Stukel (2012)

The aim of this study was to test the relationship between the quality of health-related information in press releases and news. A sample of 343 news articles reporting on the findings of research reported in 68 press releases issued by medical journals was collected. Each article was assessed by two independent raters based on whether articles contained or omitted: 1) quantification of absolute risks, 2) harms of the interventions, and 3) limitations of the research. These aspects are important to provide an adequate assessment of the quality of new health findings, but have previously been found to be absent from press releases (Kuriya, Schneid & Bell, 2008). This type of study is extremely time consuming to perform. Given the amount of time it takes for coders to assess each article, Schwartz, Woloshin, Andrews, and Stukel (2012) had to take a stratified sample of less than half of the news stories that were actually available.

The average number of news articles to contain each of the quality features was compared for press releases that contained the features, omitted the features, or for research findings that were not published in a press release. The findings, displayed in table 1.1 showed that news was more likely to contain information

about absolute risks, harms, and limitations when the press release did, versus when it did not contain the same information. Given the retrospective observational nature of this research, it cannot be concluded that the quality of press releases (as measured by presence or absence of the tested characteristics) actually causes news to be exaggerated because this is merely an association. Just because a news article reports content related to that present in the press release, it does not necessarily mean that the information was directly sourced from the press release. However, even in the absence of a confirmed causal link this research should provide impetus to improve quality of press release content if that content ends up in the news.

Table 1.1.

Data taken from Schwartz, Woloshin, Andrews and Stukel (2012) showing the percentage of news articles containing important characteristics about health research for press releases that contained or omitted the same characteristics. The difference between percentages of news for press release conditions was significant across all quality measures.

Quality measure	Presence in press release (PR)	% of news adequately reporting measure
Absolute risks	Present	53
	Absent	9
Harms	Present	68
	Absent	24
Limitations	Present	48
	Absent	16

1.8.2. Sumner et al., (2014)

1.8.2.1. Justification

Schwartz, Woloshin, Andrews, and Stukel (2012) demonstrated that the quality of press releases is related to the subsequent quality of news, but for the specific nuances of medical findings. The increase in news quality in this regard would be a positive outcome, but in its own right. In other words, a reader who is presented with more accurate information about the nuances of scientific

findings may be no better informed, and have no increased positive attitude towards science than if the information was inaccurate (Sturgis & Allum, 2004). An outcome of high importance would be the behavioural outcome related to information presented in the news, such as the examples given previously in this introduction.

Sumner et al., (2014) performed retrospective observational study, similar to Schwartz, Woloshin, Andrews, and Stukel (2012), but instead focused on the three categories of accuracy that were deemed to be the most important regarding behavioural implications for the reader: statements of relationship between variables, human inference from non human research, and advice to the reader. All three categories were framed in terms of exaggeration in press releases and news related to the content of the journal article, so the magnitude of exaggerated information relevant to human behaviour could be tracked in the transfer from press releases to news.

1.8.2.2. Exaggerated statements of relationship

The inverted pyramid writing style of news articles means that the conclusions of health-related findings, often relationships between variables, are printed typically in the first few lines of the main body of text, or even the headlines. The relationship statement is therefore the primary piece of information that readers see when they read news articles. Given that the average visit time to US online newspapers was around 2.4 minutes in 2017 (Pew Research Centre, 2018), it is likely that the statements of relationship in news articles are read more than any other aspect. As discussed previously, it is the headline statements

of news (such as “sausages cause cancer”), which are likely to alter behaviour. In the two weeks after news reports of the World Health Organisation report that processed meats cause cancer, sales of bacon and sausages had dramatically reduced (IRI, 2015). Exaggerated statements of relationship in press releases and news were classified as those that made a stronger statement (containing a higher level of causal inference) than the source journal article.

1.8.2.3. Exaggerated inference from non-human research

In addition to analysis of statements of relationship, Sumner *et al.* (2014) also examined the reporting of animal research. Only 10% of animal studies ever make it through to a human application (Van der Worp *et al.*, 2010), so reporting animal research in the news as if it was relevant to human health could have a detrimental effect if people change their behaviour based on this information. Exaggerated reporting of animal research was operationalised as cases where articles made inferences about humans based on animal research. Since Sumner *et al.* (2014) compared what the press release and news articles said, to what the journal article said, and not what the journal article did, this means that exaggeration in the journal article would not have been detected. In other words, if the journal article made an inference about humans, but the sample of the study was mice, the press release and news would not be labelled as exaggerated if they also make human inferences. This method was seen as protective of the press officers and journalists - it did not punish them for being misled by the journal article.

1.8.2.4. Exaggerated advice

Exaggerated advice was defined as advice that was present in the news or press release that was not present in the journal article, or as any advice that was more direct. For example, if the journal article were to state ‘general practitioners should reduce patient’s reliance on calcium supplementation’, the press release would be deemed to have exaggerated if it included the more direct advice ‘patients should stop taking calcium supplementation’.

1.8.2.5. Findings

Sumner *et al.* (2014) examined 462 press releases, and 668 associated news articles for three types of exaggeration. Results showed that when press releases contained exaggeration, news articles were more likely to contain the same exaggeration than when press releases did not. For statements of relationship, when press releases did not contain exaggeration, only 18% of news articles contained exaggeration. When the press release did contain exaggeration, the news was much more likely to also contain exaggeration, at 82%. For sample inference, 10% of articles contained exaggeration in the absence of press release exaggeration, and 86% of news contained exaggerated sample inference when press releases did. For advice, 17% of articles contained exaggeration in the absence of press release exaggeration, and 58% of news contained exaggerated advice when the press release did.

1.8.2.6. Interpretation

The strength of the relationship between press releases and news is striking. On the one hand it makes sense to conclude that journalists must be practicing

‘churnalism’ (Davies, 2009) – given that exaggerations in press release content are likely to end up in news. But this should be seen as an opportunity for press offices to make amendments to their practice, rather than as a problem with journalistic practices, for journalists to resolve. From a research point of view questions are raised about the mechanism by which information in press releases is absorbed by the news. The retrospective observational nature of Sumner *et al.* (2014) means that inferences regarding the apparent transfer of information between press releases and journal articles cannot be made. Experimental research is needed give better control to understand whether news writers will pick up experimental manipulations to source material. If modifications to important study related information, such as the statement of relationship between variables, are picked up by the news, this would place a greater impetus on the press release to ensure accurate news.

1.9. Synopsis

This thesis is concerned with the relationship between press releases and news articles, and is split into five sections. Chapter two is a direct replication of the work of Sumner *et al.* (2014), the first paper conducted by the InSciOut research group, of which I am a member. This is important for two reasons. Firstly, the findings of that paper are an important justification for all of the team’s following research, including the studies reported in this thesis. Secondly, replications are becoming increasingly seen as critical for the health of science in general, and should be seen as an important undertaking for early career researchers.

Given the controversy that originally motivated the research reported by Sumner *et al.* (2014), and the subsequent high level of interest in the findings of the research, chapter three investigates whether the content of science communication changed in response to the release of the paper (Sumner *et al.*, 2014).

The following three chapters move to experimental methodology to investigate the potential effects of press release content on subsequent news article content and selection. Chapter four reports my contribution to a multi-year, multicentre, randomised controlled trial in which we manipulated the content of real-world press release prior to their publication in order to test the effects of manipulations to statements of relationship and caveats regarding study design on subsequent news coverage. This study reported in this chapter has been submitted for publication: Adams, Challenger, Bratton, Boivin, Bott, Powell, Williams, Chambers, and Sumner (manuscript submitted for publication).

The experiment in chapter five investigates the same manipulations to press releases on the selection and content of news, but in a more controlled experimental setting using masters-level journalism students. The nature of this experiment allowed for the comparison of the effect of press release content using almost identical press releases between-subjects. This study was designed to compliment the randomised controlled trial reported in chapter four. The phase of this experiment reporting on the effects of press release content on news selection is included in a paper submitted for publication: Bott, Bratton, Diaconu,

Adams, Challenger, Boivin, Williams, and Sumner (manuscript submitted for publication).

Chapter six reports on an experiment investigating the effect of writing-style instructions on the content of science articles. In this study, undergraduate participants are tasked with writing an article based on a health-related press release, but under the explicit instruction to either write accurately, in a more concise way, or in a way that is more appealing. This is to test whether the motivation for writing an article might affect its content.

Finally chapter seven discusses the thesis and its implications for the reporting of health-findings. I make the argument that there appears to be no penalty for accurate science reporting, but the potential for negative outcomes without careful practice of press relations. There is enormous potential for improvements in the science media process.

CHAPTER 2 - REPLICATION

2.1. A note on contribution

I joined the project after the data collection phase and became responsible for data handling and organisation of the coding regime, as well as completing a significant portion of the article coding. I then performed the data analysis. The InSciOut research group conceived the data collection method. Dr. Louise White (research assistant), Dr. Rachel Conde Adams (post-doctoral researcher), and Aimeé Challenger (research assistant) performed the search for press releases, journal articles, and news articles. These, in addition to Seemu Ali, Jemma Pitt, and Thomas Casey (undergraduate students) assisted me in coding the articles and performing inter-rater consensus checks. The InSciOut research group created the coding sheet used to record data; this coding sheet is a modification of that used in the previous study (Sumner *et al.*, 2014). Data from Sumner *et al.* (2014) are presented for comparison throughout this chapter and are indicated appropriately. Dr. Geoffrey Megardon (post-doctoral researcher) assisted with data handling by creating an SQL database.

2.2. Introduction

The following study is an attempt at replication of the findings of the study by Sumner *et al.* (2014). Outlined below are the events that became the impetus for the research.

Between August 6th, and August 11th 2011, riots were taking place across the United Kingdom (Rogers, Sedghi, & Evans, 2011). The first incident took place in Tottenham on Saturday 6th August, when protests regarding the killing of

Mark Duggan by police on Thursday 4th August turned violent. Over the next few days, numerous riots occurred across the country in which thousands of individuals took part in vandalism, looting, and violence, resulting in the mass deployment of police.

Just prior to the riots, researchers at Cardiff University, The Johns Hopkins University, the Kennedy Krieger, and University College London published a study examining the relationship between γ -amino butyric acid (GABA) concentration in the dorsolateral prefrontal cortex, and the urgency facet of impulsivity in a sample of males (Boy *et al.*, 2011). The study described a correlation between the concentration of GABA and variation in the urgency trait, where lower concentrations of GABA were associated with higher scores on measures of urgency. The conclusions were that at most, this finding could help to clarify the relationship between GABA and psychiatric disorders that are characterised by certain cognitive symptoms. Despite this simple finding, subsequent media coverage included claims that were not present in the research paper. For example, the newswire service of the Press Association released an article entitled “Brain chemical lack ‘spurs rioting’”; a headline that could be interpreted to imply that GABA concentration in the dorsolateral prefrontal cortex caused individuals to participate in the riots. The use of quotes around the term ‘spurs rioting’ implies that previous authors have used the term, but there was no reference to rioting in the original research article. The Daily Mail ran the headline “Rioters have 'lower levels' of brain chemical that keeps impulsive behaviour under control”; again, relating the research finding to the behaviour of rioters. The headline published by The Sun reported on a “Nose spray to stop

drunks and brawls”; a statement devoid of the variables actually reported in the study, and adding a further layer to the findings implying that the potential negative effects of impulsivity could be cured by administering a drug.

The three news articles mentioned above were retracted or changed, but at the time of the riots were discussed in a Guardian article by two of the authors of the original research, and a colleague (Sumner, Boy, & Chambers, 2011). The article titled “Riot control: How can we stop newspapers distorting science?” did not entirely lay blame on journalists, but it strongly raised concerns about the distortion of science in the media. Specifically that the already compromised Press Association newswire story appeared to act as the main source of information for further news articles, meaning that the article’s inaccuracies were proliferated and exacerbated by the “zombie-like repackaging” (Sumner, Boy, & Chambers, 2011) of information by journalists. The article also laid some blame on journalists and editors for the lack of regard for the reputation of scientists and the lack of regard for public understanding. This prompted discourse amongst stakeholders in the science news process, which led to a debate at The Royal Institution (The Royal Institution, 2012). The debate was intended to act as a stage for the deliberation, from both the scientists’ and the journalists’ points of view, of practical steps to take towards improved communications (Scott, 2012).

The emerging story is that much of the blame for exaggeration attributed to journalists by scientists, and the blame for poor communication skills attributed to scientists by journalists had overlooked the intermediary step in their

communication – the press release. Williams and Clifford (2009) conducted a study canvassing the opinions and experiences of 89 current and former journalists, and five senior news editors. Their survey and focus group data revealed that the modern pressures of the journalistic environment, such as increased workloads due to multiplatform news production, the mandatory requirement for covering an increasing calendar of perennial topics, and the battle to not be left out by missing a good story that a competitor publishes (so called “pack-journalism”), leaves less time for core journalistic work, like fact checking and independently investigating new stories and alternate sources. Such pressures mean that the press release is seen as an increasingly attractive resource for journalists looking to quickly produce content, since press releases are already presented in a more digestible format for the general reader, and are emailed to journalists daily. Davies (2009) suggested that this environment has fostered the “churnalism” behaviour, or the heavy reliance on press releases, to the extent that science news articles are largely reproductions of the content already provided by press officers. This was based on the finding that the content of around 40% of news stories reporting on health and nature are entirely or mainly reliant on press releases (Lewis, Williams, Franklin, Thomas, & Mosdell, 2006).

Researchers from Cardiff University involved in the debate at the Royal Institution, along with other colleagues from Cardiff University formed the multidisciplinary InSciOut research group to investigate the relationship between press releases and news. Their focus was on the three exaggeration types described in chapter 1: the introduction of new advice, or the inclusion of more direct advice than the journal article; the use of a stronger statement of

relationship between the two variables reported on in the journal article; and inference regarding humans based on non-human research. Using data from press releases published in 2011, Sumner *et al.* (2014) examined the presence of all three types of exaggeration in press releases and news articles using the journal article as the baseline. News articles reporting on exaggerated press releases were found to contain the same type of exaggeration more frequently than news articles reporting on press releases that were representative of the journal articles (in other words, press releases that did not contain exaggeration). Furthermore, exaggeration in press releases was not associated with an increased likelihood for news to report on the story – nullifying the potential argument that exaggeration would make press releases more appealing in the e-mail inboxes of journalists, and be more likely to be turned into a news story. Though observational, these findings contributed to the notion that the press release is an important part of the science communication trajectory. This was not the first study to demonstrate a relationship between news content and press release content. Previously, Schwartz, Woloshin, Andrews, and Stukel (2012) demonstrated that when absolute risk, harms relating to interventions, or research limitations were presented in the press releases from major medical journals, the news was more likely to report the same important facets of the research. As a further step, when no press release was produced, news was more likely to include this important information than when press releases were produced but did not contain such information. This suggests that poor quality press releases could be worse for news than no press release being issued.

Both of the mentioned studies, whilst contributing strong findings, were based on retrospective observational designs. There is a need for experimental data to conclude whether press releases have a significant effect on news articles when other factors are accounted for. The InSciOut team devised a randomised controlled trial to discover whether interventions in the science communication cycle could affect news content. In order to take a baseline measurement of the exaggeration present in science communication just prior to the commencement of the trial, health-related articles were sampled from two time-points: the period of January to June 2014, and January to June 2015. Given the discourse outlined above, and the level of discourse following the release of the results of Sumner *et al.*, (2014) (as identified by the article's high Altmetric score), this dataset also enables an analysis of the potential impact of the findings of Sumner *et al.* (2014). This existence of this dataset provides a unique opportunity to attempt a replication of the findings of Sumner *et al.* (2014).

Schat, Bossema, Numans, Smeets, and Burger (2018) published a smaller scale replication of the Sumner *et al.* (2014) using 129 health-related press releases from Dutch universities and 185 associated news articles collected in 2015. The relationship between exaggeration in press releases, and exaggeration in news was still present in the data, with higher levels of exaggeration present in news articles reporting on exaggerated press releases. Contrary to the findings of Sumner *et al.* (2014), Schat *et al.* (2018) found that press releases with exaggeration were more likely to be picked up by the news than press releases that did not contain exaggeration.

The present study aimed to attempt replication of the findings of Sumner *et al.* (2014) using data collected retrospectively. It was not expected that there would be a notable disparity between the 2011 and 2014/2015 data other than potentially with regard to levels of exaggeration in human inference from non-human samples (explained further below). The dataset is comprised of press releases from the 20 institutions that were members of the Russell Group in 2011. Between 2011 and 2014 (the start of the time period for data used in this replication) the majority of the institutions in this sample were involved in the Declaration on Openness on Animal Research (2012), which was a commitment to the creation of the Concordat on Openness on Animal Research in the UK (2014). This movement bound its signatories to a number of commitments including specifically to “include information about that animal research in relevant communications, including media releases” (Concordat on Openness on Animal Research in the UK, 2014). This explicit commitment was signed by all but two of the institutions in the sample used in this replication; one of the unsigned institutions does not have a heavy scientific research focus, and is unlikely to contribute many press releases to the sample. It is expected that the inclusion of unwarranted human inference in press releases would be diminished between the 2011 dataset of Sumner *et al.* (2014), and the 2014/15 replication dataset. For the relationship between exaggeration in press releases and the presence of news stories, it is unclear whether the relationship reported by Schat *et al.* (2018), is a product of factors related to the Dutch sample (in which case, the present study would be expected to show no relationship between press release exaggeration and uptake by the news), or whether this represents a shift in news behaviour towards favouring more strongly worded press releases (in

which case, exaggerated press releases would be expected to be related to more frequent news uptake).

2.3. Method

2.3.1. Data collection

Press releases from 20 Russell Group universities were used as participants in this study, the same sample as Sumner *et al.* (2014). The Russell Group is a group of 24 prominent institutions from the United Kingdom, all with a notable high standard of research activity. For example, in the 2014 Research Excellence Framework, the Russell Group accounted for 68% of all four-star rated research in the country (Russell Group, 2014). Since the data collection by Sumner *et al.* (2014) took place in 2011, when there were only 20 members of the Russell Group, the four institutions that joined the Russell Group in 2012 (Durham University, University of Exeter, Queen Mary University of London, and University of York) have been excluded from this replication.

The sample period was January to June 2014, and January to June 2015. This time frame was selected in order to make the comparisons in chapter 3, in which I examine the change in journal article, press release, and news characteristics over time. For the analysis in this chapter, data from both periods was combined into a single dataset. Online repositories (the universities' websites, and EurekaAlert.org) were searched for any press releases from the included institutions. This resulted in a corpus of 4476 press releases. For each institution, the number of available press releases varied considerably, with the lowest output being 90 press releases, and the highest output being 517 press releases; a difference of 427 press releases over the same time-period. The sample was

restricted to those relevant to human-health (mainly those reporting on psychological and biomedical findings), and whether they reported on a single, published, peer reviewed research article. This left 890 relevant press releases. The lowest number of relevant press releases from an institution was seven, and the greatest number of press releases from a single institution was 111. In order to reduce variation in number of relevant press releases between the institutions, and therefore reduce any bias in the findings towards the behaviour of specific universities, a cap of 10 press releases for each time period for each institution was applied, leaving up to 20 press releases per institution. This was achieved through a randomisation process. This resulted in a sample of 351 press releases (see figure 2.1).

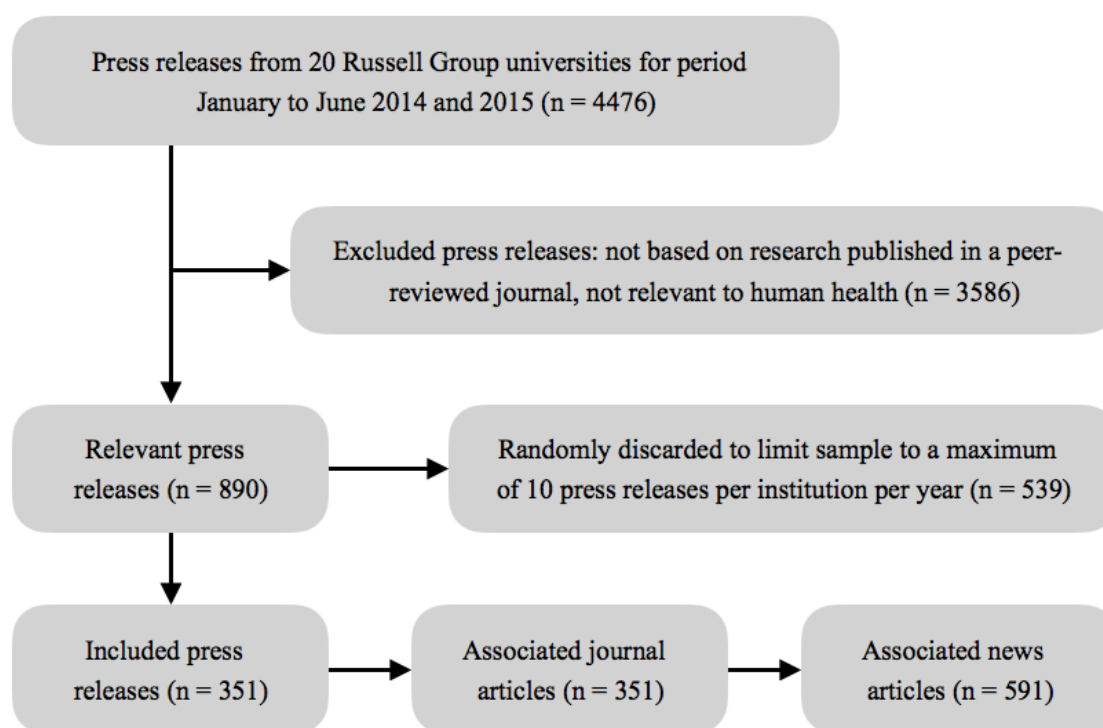


Figure 2.1. Press release collection process.

For each selected press release, relevant news articles were collected via keyword searches using Google Search, and the Nexis database (LexisNexis,

New York, NY). The search was conducted for relevant news articles (i.e. those which make reference to the source research) up to 28 days after publication of the press release, and up to one week before. Extending the search to one week before was to account for the possibility that a news story was released before the embargo was lifted.

2.3.2. Article Coding

Each press release, associated journal article, and any related news articles (herein referred to as article set), were searched for a number of attributes, which were recorded using the coding sheet included in Appendix 2.1. This coding sheet is a modification of that used by Sumner *et al.* (2014) and acts as a rigid framework for interpretation of the attributes of interest in articles, such that two independent coders should generally come to a consensus. The original coding sheet was designed to comprehensively record as much information about the article sets as possible, such that the resulting data set may be used for multiple future studies. However, for the studies outlined in this thesis, my research group employed a strategy of reducing the data collected to the minimum required to allow for the comparisons in the studies outlined in this thesis. This means that the average time taken for a single coder to code an article set reduced from around three hours for the previous coding sheet (Sumner *et al.* 2014), to less than an hour. Given the amount of time it takes for such a coding task, the coding sheet was created to record data for a number of different studies. Information to be recorded in blue cells was used for the calculation of data used in the studies in chapter two and three. Data recorded in red cells was used specifically for the randomised controlled trial outlined in chapter four. Information in the red 'Press

Release Characteristics' section was recorded to assist with the trial procedure, and to record information that press offices requested we record as a condition of their participation. Specifically, 'expected level of news uptake', and 'particular desired outcome' were added to record the press office's expectations about number of news articles, and any other free-text expectations, so that they can be compared to actual outcomes at the end of the study. For each article set, two researchers independently recorded the attributes of each article, and subsequently both researchers' work was compared electronically for disagreements. The coders then met to discuss disagreements and come to a consensus. Though I did not collect data regarding the types of disagreements, from my own experience as a coder using this protocol, disagreements were generally simple to resolve. For example, issues caused by accidental keystrokes, pieces of information overlooked when reading articles, or information inputted into the wrong cells of the coding sheet. In an attempt to avoid the possibility that one coder could influence the other when coming to agreements, if a particular disagreement could not be resolved easily, a third researcher not involved in the coding of that article set was consulted to make a decision between the two alternative interpretations; the third coder's decision was final. Although Sumner *et al.* (2014) demonstrated with simulations that a 10% rate of disagreement between coders would not influence their conclusions, the present study intended to avoid the possibility of such an impact by avoiding variation by using the mentioned coding and consensus arrangement.

The comparisons of interest were the same as those in Sumner *et al.* (2014); the presence and strength of advice, the strength of the statements of relationship, and the type of sample mentioned.

2.3.2.1. Coding of advice

For advice there were three levels coded (see table 2.1). Advice was coded if it appeared anywhere in the title, abstract, or main text of the journal article, and in the titles or main text of press releases and news articles. If multiple pieces of advice appeared in an article, the strongest piece of advice was recorded. An article was deemed to have exaggerated if it contained a higher code than its article of comparison. For example, ‘explicit advice to the reader’ in a press release is an exaggeration of ‘explicit advice, not to the reader’, in a journal article. With this method, a piece of advice for a medical practitioner to prescribe a different drug for certain patients is not seen to be as strong as a piece of advice for patients themselves to change their behavior. Further, either type of advice, in a press release, would be considered an exaggeration of an absence of advice in a journal article. Cases where neither article contained any advice were excluded from the analysis.

2.3.2.2. Coding of statements of relationship

Statements of relationship were coded on a seven-point scale (table 2.1) and were recorded from the abstracts and discussion sections of journal articles; the strongest statement was selected for use in the analysis. The ‘cause’ category included any relationship that was deemed to describe one variable directly affecting another, including phrases such as modifies, changes, and impacts.


'Can cause' included any causal statement with the word 'can' to modulate the strength of the relationship to suggest that one variable has the power to affect another, but does not always do so. 'Conditional cause' statements, like 'could cause', or 'might cause', are those that suggest uncertainty, but potential causation. Ambiguous statements, such as 'linked', and 'connected' are those that provide more information than correlational statements, but do not suggest causation. Correlational statements are any that suggest that changes in variables are related, but the statement makes no inference about the influence of one on another. Statements that explicitly state that there is no relationship were classified as 'does not cause'. Finally, 'no cause mentioned' was used in cases where there was no other statement of relationship.

For press releases and news articles, the strongest statement from the first two sentences, that were not context, from the main body of text were used. This strategy was employed because of the inverted pyramid structure of news where the most pertinent points are presented first. Only article sets where the journal article reported on observational cross sectional, or observational correlational designs were used for analysis.

2.3.2.3. Coding of samples

Samples were rated on a three-point scale (table 2.1), and only journal articles with non-human samples were used for the comparison.

Table 2.1.
Coding categories for advice, language, and sample. Categories are listed with descending strength, with categories higher up the table trumping categories below.

Strength	Advice	Language	Sample
	Explicit advice to reader	Cause	Human
	Explicit advice not to reader	Can cause	Implicitly human
	No advice	Conditional cause	Non-human
		Ambiguous	
		Associative	
		Does not cause	
		No cause mentioned	

2.3.3. Analysis

Three types of analysis were carried out: quantifying levels of press release exaggeration, news exaggeration relative to press release exaggeration, and news uptake (whether news reported on each press release/journal article pair) relative to press release exaggeration. Generalised estimating equations (GEE) were used to provide percentage exaggeration rates and 95% confidence intervals.

Exchangeable working correlation was used to adjust for the clustering of multiple articles to one source (for example, multiple news articles to one press release). This type of working correlation matrix assumes that data from different articles arising from the same source are equally correlated; in this data set we have no reason to assume that different articles from the same source would be differentially correlated with each other. For press release characteristics and the calculation of news uptake rates, GEE and exchangeable working correlation were used to adjust for the clustering of multiple press releases to each institution. In the analysis of news exaggeration, GEE was used to adjust for the clustering of multiple news articles to any single press release. A binary

distribution was specified since the outcome variables in each comparison are the presence or absence of exaggeration, or presence or absence of news. A logit linking function was employed to allow for the easy interpretation of odds ratios. For the interpretation of odds ratios, in the news uptake and news exaggeration analyses, exaggerated press releases were treated as the intervention, and representative press releases were treated as the control. Odds ratios greater than 1 are indicative of a relationship between press release exaggeration and the outcome variable.

2.3.3.1. Advice

Unlike with comparisons of statement of relationship and the type of sample, the journal article does not contain an attribute for the baseline comparison of advice. That is, the sample mentioned in any article can be compared to the actual sample used in the journal article; or the strength of relationship can be compared to the maximum strength implied by the study design (i.e. associative statements from correlational designs); in comparison, there is no attribute of the journal article to compare to advice in press releases and news articles, other than whether advice was given by the journal article authors. So to compare the exaggeration of advice across the articles, only cases where advice appeared in at least one article in the article set were selected. This means that cases where no advice appeared in either the journal article, press release, or news article(s) for an article set were excluded, meaning that such cases would not be treated as a lack of exaggeration; rather, they were treated as cases where exaggeration could not be measured. For the measurement of press release exaggeration rates, a total of 74 press release/journal article pairs contained advice. For the comparison of

advice in news articles, there were 70 press releases and journal articles, with 248 associated news articles, where at least one in each article set contained advice.

2.3.3.2. Statements of relationship

This analysis only included cases where the design of the study reported in the journal article was observational. As with the previous research (Sumner *et al.*, 2014) I only selected cases of observational longitudinal, and observational cross-sectional design. Interventions, computer models/simulations, qualitative designs, and meta-analyses of observational studies were all coded for, but since it is open to debate whether causal inference can be made from these designs, they are excluded from this analyses. In total there were 154 press releases available for analysis. For the calculation of exaggeration rates in news, 58 of these press releases had news. There were 237 associated news articles.

2.3.3.3. Inference from non-human samples

Studies reporting on human samples were excluded from these analyses such that any article with conclusions that make inference about humans could be deemed to contain exaggeration. There were 117 press releases based on studies with non-human samples that could be used to calculate press release exaggeration percentages. For the analysis of exaggeration present in the news, there were 38 press releases available with 129 associated news articles.

2.4. Results

2.4.1. Press release exaggeration

For the comparison of the presence of advice in press releases in excess of that present in the journal articles, 51% (95% confidence interval = 40% to 62%) of press releases contained advice not present in the journal article, or advice that was more direct than that in the journal article. For the comparison of the strength of statements of relationship, 27% (95% confidence interval = 21% to 35%) of press releases contained a statement that was more strongly worded than the strongest statement present in the associated journal article. For the comparison of human inference from non-human samples, 21% (95% confidence interval = 15% to 30%) of press releases contained implicit or explicit references to human samples when the journal article did not. See figure 2.2 for a comparison of press release exaggeration in this data and Sumner *et al.* (2014).

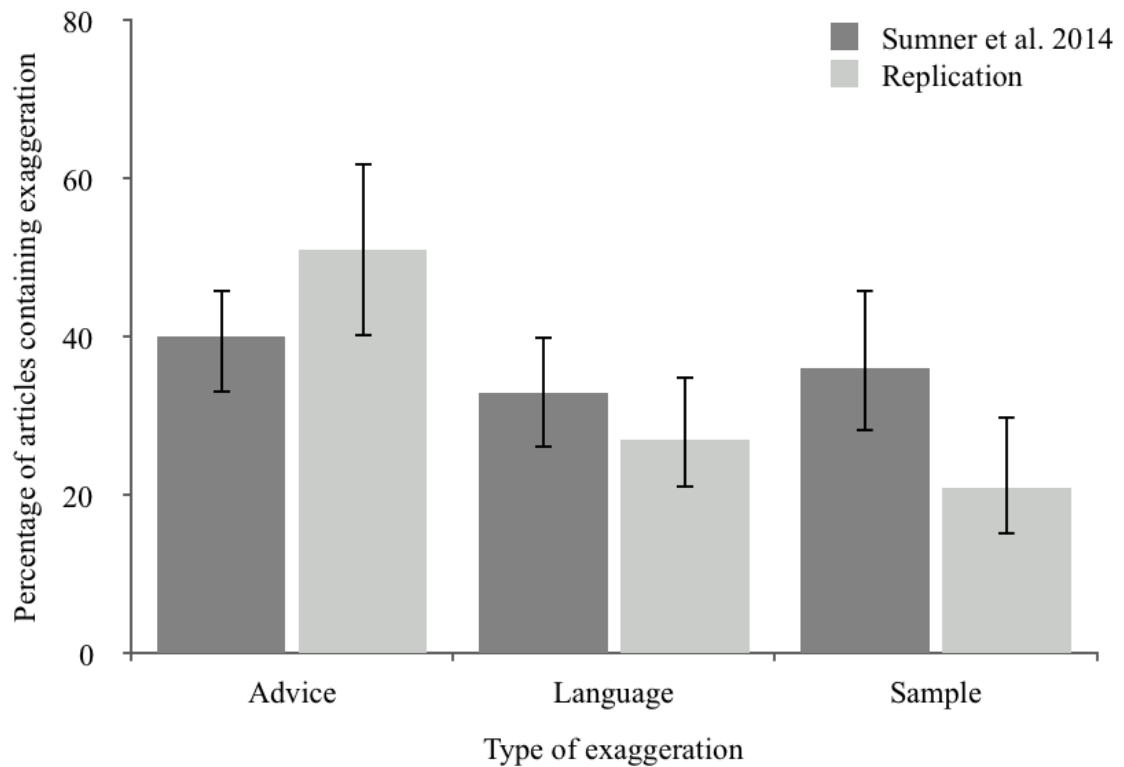


Figure 2.2. Percentage of press releases containing exaggeration for each of the three categories of exaggeration (light gray bars). Data from the same analyses by Sumner et al. (2014) are presented for comparison (dark gray bars). Error bars are 95% confidence intervals.

2.4.2. Exaggeration in news relative to exaggeration in press releases

Overall, 55% (95% confidence interval = 44% to 65%) of news articles contained new advice, or a higher level of advice than the associated journal article. When press releases contained exaggerated advice, 49% (95% confidence interval = 34% to 65%) of the related news reports were also exaggerated in the same way. Conversely, when the level of advice in the press release was not in excess of that found in the journal article, 60% (95% confidence interval = 46% to 72%) of the associated news articles contained exaggerated advice. There was no relationship between exaggerated press releases and the presence of exaggerated advice in the news (difference = 11%, 95% confidence interval = -9.9% to 31.9%; odds ratio = .7, 95% confidence interval = .3 to 1.5). This

contrasting outcome to the finding of Sumner *et al.* (2014) is driven by the relatively high number of exaggerated news articles in article sets containing representative press releases. Figure 2.3 demonstrates this outcome compared to the data from Sumner *et al.* (2014).

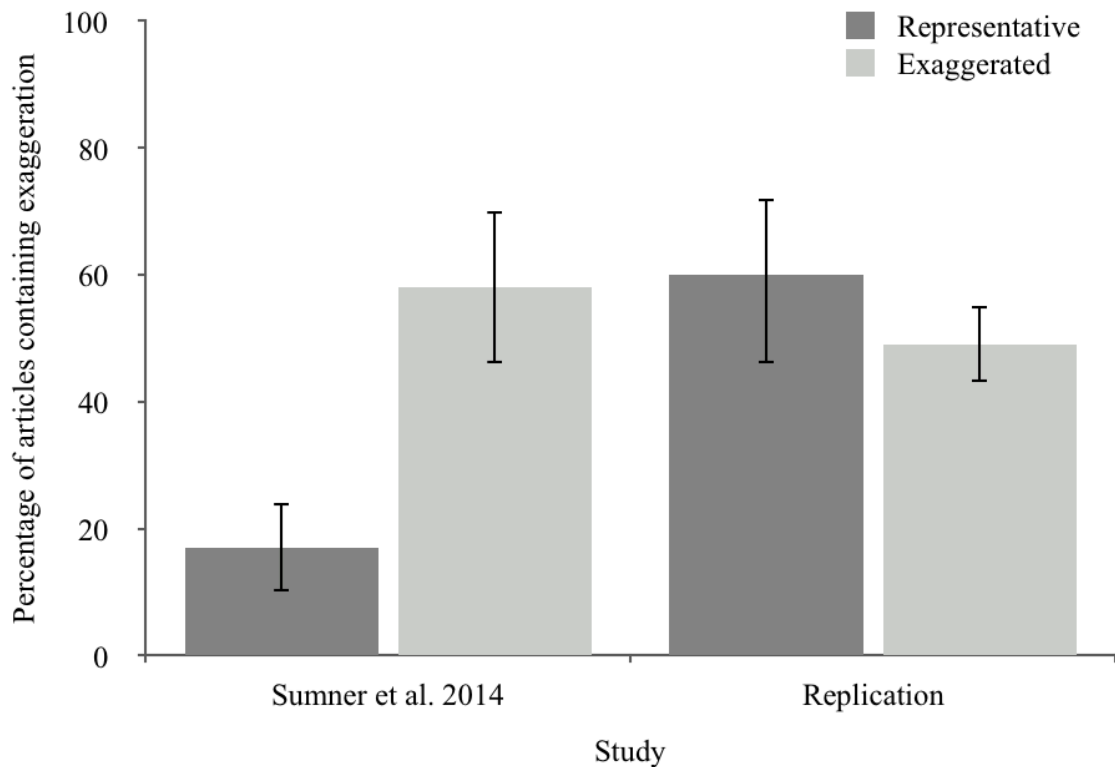


Figure 2.3. Percentage of news articles containing exaggerated advice in article sets containing representative (dark gray bars) and exaggerated (light gray bars) press releases. The data for the same comparison by Sumner *et al.* (2014) are displayed for comparison.

For the comparison of the strength of statements of relationship, the language used was more deterministic than that present in the journal article in 49% (95% confidence interval = 37% to 61%) of news articles. When press releases contained exaggeration of the language used in their related journal articles, 82% (95% confidence interval = 68% to 91%) of the associated news articles also contained exaggeration compared to 16% (95% confidence interval = 10% to

26%) of news articles when the press releases were not exaggerated. The difference between rates of exaggeration between conditions was 66% (95% confidence interval = 52.2% to 79.8%), and the odds of exaggeration in news were 23.7 times higher in relation to exaggerated press releases than representative press releases (odds ratio = 23.7, 95% confidence interval = 9.0 to 62.2). For a comparison of this data to the same analysis from Sumner *et al.* (2014), see figure 2.4.

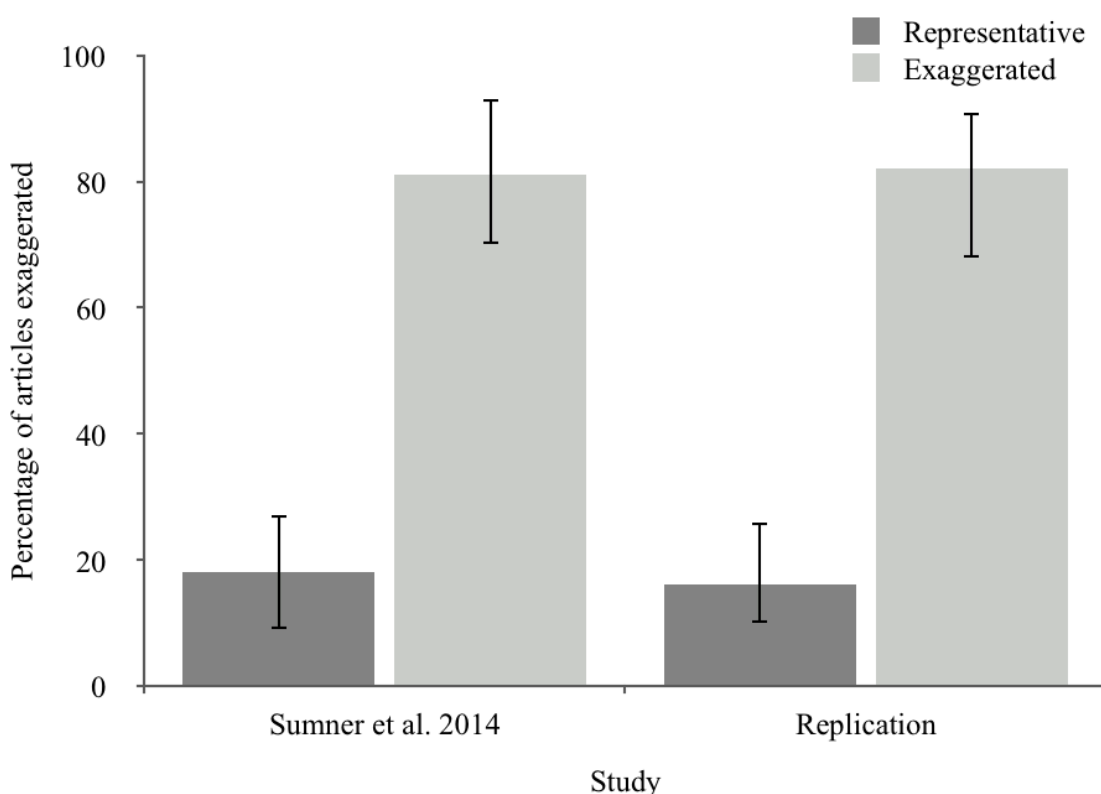


Figure 2.4. Percentage of news articles containing exaggerated statements of relationship in article sets that contain representative (dark gray bars) or exaggerated press releases (light gray bars). Data from Sumner *et al.* (2014) are presented for comparison.

For the comparison of reported samples, 33% of news articles included statements that made inferences relating to humans in excess of those present in the news articles. When press releases contained exaggerated statements, 72% (95% confidence interval = 46% to 88%) of the related news contained

exaggeration, compared to 9% (95% confidence interval = 3% to 21%). The difference between conditions was 63% (95% confidence interval = 39.4% to 86.6%) and the odds of exaggeration in news were 26.5 times higher (95% confidence interval = 6.1 to 116.0). See figure 2.5 for a comparison of this data to the same analysis by Sumner *et al.* (2014).

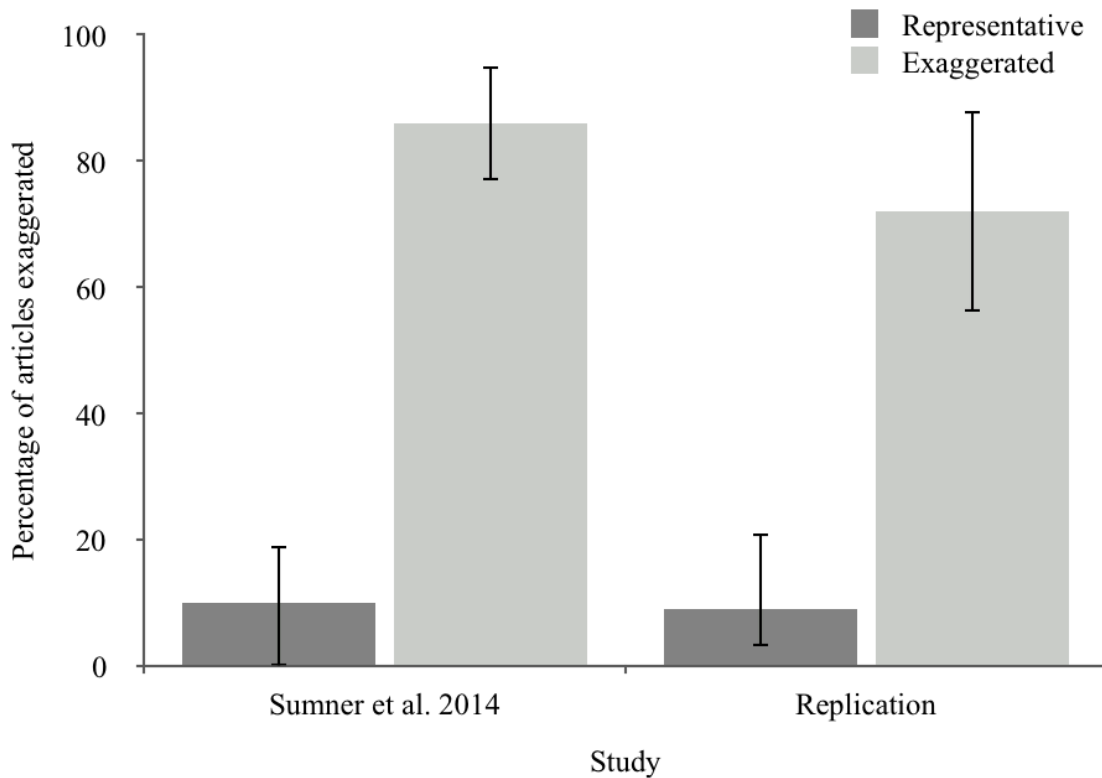


Figure 2.5. Percentage of news articles containing human inference from non-human samples in article sets that contain representative (dark gray bars) or exaggerated press releases (light gray bars). Data from Sumner *et al.* (2014) are presented for comparison.

2.4.3. News uptake relative to exaggeration in press releases

There was no relationship between the presence of news coverage for press releases that contained exaggeration versus those that did not in any of the three analyses. For press releases with exaggerated advice, 57% (95% confidence interval = 39% to 74%) had associated news stories compared to 45% (95% confidence interval = 31% to 59%) for press releases with representative levels

of advice (12% difference, 95% confidence interval = -11.5% to 35.5%). For press releases with exaggerated statements of relationship, 57% (95% confidence interval = 40% to 72%) had associated news articles compared to 56% (95% confidence interval = 44% to 66%) for press releases with representative statements of relationship (1% difference, -18.6% to 20.6%). For press releases with human inference from non-human samples, 44% (95% confidence interval = 27% to 62%) had associated news compared to 35% (95% confidence interval = 26% to 46%) for press releases mentioning appropriate samples (9% difference, 95% confidence interval = -12.1% to 30.1%). Table 2.2 provides a comparison of the level of uptake for each analysis compared with the equivalent results from Sumner *et al.* (2014).

Table 2.2.

Percentage of article sets containing news stories compared between Sumner et al. (2014) and this replication.

		Sumner <i>et al.</i> (2014)		Replication	
		Percentage with news	95% confidence interval	Percentage with news	95% confidence interval
Advice	Representative	52	43 - 60	45	31 - 59
	Exaggerated	59	48 - 69	57	39 - 74
	Difference	7	-6 - 21	12	-12 - 36
Language	Representative	50	41 - 60	56	46 - 66
	Exaggerated	57	44 - 69	57	40 - 72
	Difference	7	-9 to 22	1	-19 - 21
Sample	Representative	43	32 - 56	34	26 - 46
	Exaggerated	50	35 - 66	44	27 - 62
	Difference	7	-13 - 27	9	-12 - 30

2.5. Discussion

The purpose of this study was to attempt a replication of the main findings of Sumner *et al.* (2014). The previous research examined the presence of advice, statements of relationship, and the sample mentioned in journal articles, press releases, and news articles. It found that exaggeration of these traits in news articles was related to exaggeration in news, such that news articles were more likely to contain exaggeration if the associated press release was exaggerated. Their data did not support the notion that exaggerated press releases would be more likely to have their stories taken up by the news. In comparison, the replication analyses described in this chapter largely support findings of Sumner *et al.* (2014), with some slight exceptions. For press releases, exaggeration in all three categories was approximately equal to the levels in Sumner *et al.* (2014), but with advice tending towards being more exaggerated, and human inference from non-human samples tending towards being more representative. In this replication, the relationship between exaggeration in the news, and the same type of exaggeration in press releases was similar for all three exaggeration types, except in the case of exaggerated advice in the news. Exaggeration in news articles was elevated even when the associated press releases contained representative advice, and this level of exaggeration in the news was elevated beyond the level reported in Sumner *et al.* (2014). Consistent with the Sumner *et al.* (2014), levels of news uptake were approximately equal between representative and exaggerated press releases across all three exaggeration types. There appears to be a slight elevation in news uptake for press releases with exaggerated advice, and sample, but this was not found to be a strong enough effect.

2.5.1. Differences to the Sumner et al. (2014)

The trend towards lower levels of human inference from non-human findings in press releases may be indicative of the success of the Concordat on Openness on Animal Research in the UK. The concordat was signed by the majority of the institutions in this replication from May 2014 onwards (i.e. during the sample period), and was developed in the Declaration of Openness on Animal Research from the end of 2012 onwards by many of the sample institutions; that is, it was conceived after the publication of the press releases sampled from 2011 by Sumner *et al.* (2014). This heavy focus on enhancing communications with the media with regard to animal research included the commitment that each signatory would “include information about that animal research in relevant communications, including media releases” (commitment 2 of the concordat, Concordat on Openness on Animal Research in the UK, 2014). It would seem likely that if there is a genuine reduction in inaccuracies in reporting of animal research, the concordat is likely to be responsible, since the majority of the institutions sampled explicitly agreed to this approach. The proportion of news articles containing unwarranted human-inference, as a function of sample inference type in press releases, did not change, but the overall proportion of press releases and news articles exaggerating tended towards lower levels than Sumner *et al.* (2014). It would seem likely that this pattern could be indicative of a reduction in exaggeration making it into news articles merely because journalists were less frequently given the opportunity to be exposed to exaggeration in press releases, but this notion would need testing in an experimental context. Nonetheless, this should be taken as an example of the potential positive outcome of institutions uniting to address the issues with

science communication. If similar support can be gained within the science communication community for initiatives such as the Academy of Medical Sciences' press release labeling system (Academy of Medical Sciences, 2018) or the Science Media Centre's best practice guidelines for journalists (Science Media Centre, 2012), the types of exaggeration tested here, and mentioned in those guidelines, might be diminished. Both of these publications mention being clear about the nature of the sample, and both provide suggestions for how to handle statements of relationship, but the extent of the proposal for how to handle advice is limited to: "Distinguish between findings and interpretation or extrapolation; don't suggest health advice if none has been offered" (Science Media Centre, 2012). Although this suggestion, that reporters should only print advice if advice is presented to them, with the press release as their common information subsidy (Nelkin, 1995), would not remedy the issue of exaggerated advice in press releases being picked up by the news. The evidence in this study suggests that journalists, and potentially press officers, have increasingly begun to extrapolate advice in recent years, to the point that there is no relationship between misstated advice in press releases and news articles.

The potential slight elevation in exaggeration of advice in press releases, and in the elevation in news that are based on representative press releases, may be an underestimation of the actual levels of exaggeration in comparison to the Sumner *et al.* (2014). In addition to the three levels of advice coding used in this study, Sumner *et al.* (2014) also included a fourth category, "implicit advice" between the 'no advice', and the 'explicit advice but not to the reader' categories. The implicit advice category presented a lot of ambiguity in interpretation between

coders. For example, although a phrase such as “a daily 30 minute walk with a friend or family member might be a good way to remain socially and physically active” could be interpreted to be suggestive of a change of behavior, rather than a recommendation it could equally just be interpreted as a mere postulation, with no implication. It is likely that the inclusion of instances of implicit advice in the coding protocol would push levels of advice towards the upper confidence intervals. Whether implicit advice would be effective in altering readers’ behavior is open to debate, but given the extensive scale of the interest in online health advice (Fox, 2006) it would be safe to assume that even a small reduction in inaccurate advice could have an effect. This extensive interest in online health advice; the high self-reported levels of impact of online health advice on peoples’ behaviour (Pew Research Centre, 2009); and the increasing proportion of news companies’ advertising revenue being generated via their online businesses (Barthel, 2017), may combine to passively give rise to this potential trend for increased exaggeration in advice. News businesses generate their revenue through sales of newspapers and advertising revenue generated through both online and offline stories, so if customers’ money and clicks are increasingly attracted to articles with advice, this could reinforce the publication of those articles.

2.5.2. Limitations

As alluded to in the methods section, a potential problem arises in the definition of exaggeration in advice – what is the objective justification for advice? For both other types of exaggeration, a characteristic of the study can be used as justification for categorisation. For the use of human inference from non-human

samples, the actual sample of the research can be compared to the sample stated by an article; if the sample is mice, but an article provides an explicit inference about humans, this is an exaggeration. For exaggeration of statements of relationship, the study design can be compared to the language used when referring to the relationship between variables in an article; so if the study design is observational, but an article makes a causal claim, this is an exaggeration. In this regard, the sample mentioned and relationship stated in journal articles could be analysed for exaggeration in comparison to the articles' own attributes. This is not the case for advice, where exaggeration is defined as being the presence of advice in an article (press release or news article) in the absence of advice in the journal article, or the presence of more direct advice than that in the journal article. With no objective basis, the formulation of advice would be merely down to the writer's own interpretation. In this case the researcher could be seen as a better authority for whether advice should be given, since researchers have more extensive experience in their research areas than science communicators further along the chain, and this could justify the Science Media Centre's recommendation to journalists to not print health advice if none has been offered – as long as the press officers follow the same recommendation. Another alternative might be to calculate advice exaggeration with study design as a comparison, since evidence for causation (for example, randomised control trials) could be seen as providing better justification for behavioural change than correlational evidence (such as from observational cross sectional studies). But it is unlikely that press offices and news outlets would commit to a stipulation whereby they would only print advice for findings which originate from

experimental studies. Indeed, many scientists would also probably disagree that advice cannot be formulated from correlational evidence.

The decision to exclude cases where neither article contained advice, and cases where an experimental design were used, was taken in order to only record instances where the source articles contained advice that could be exaggerated, or where an article spontaneously gave advice that was not present in a source article. The downside of this approach is that the overall rate of exaggeration would be calculated to be higher than if the excluded cases were included, and this could be seen as a source of bias.

Limitations to Sumner *et al.* (2014) also apply to this replication. In the previous study, Sumner *et al.* (2014) described how the retrospective observational nature of the study design could not be used to make inferences about whether exaggeration in press releases causes exaggeration in news articles, and that same limitation is present here. Though it is clear that press releases are important sources of information for journalists, there are many sources of variation, and potential influence between press releases other than those reported in this study (for example: word count, study topic, reported statistics) which may play a role in news selection and content. It is unclear to what extent the relationships reported would exist for two identical press releases that differ only in the variables of interest, but this is examined in later chapters.

In addition, this replication has some undesirable characteristics in comparison to the replication by Schat, Bossema, Numans, Smeets, and Burger (2018). In both

studies, press releases were treated as the participants, and in that regard the sample in this study and that used by Schat *et al.* (2018) are independent of those tested by Sumner *et al.* (2014). But the sample of Dutch language articles used by Schat *et al.* (2018) represents a better opportunity to evaluate the robustness of the results of Sumner *et al.* (2014) in a wider European context. Since the replication reported in this thesis examined articles from the same press offices, and the same newsrooms, this study is a more controlled replication. It would be expected *a priori* that the Dutch sample would generate more possible variance, and indeed this is reflected in the finding that the odds of news uptake were higher for exaggerated versus non-exaggerated press releases. Also, whereas the team of researchers who conducted the Dutch replication were independent of the authors of the original study, the data collection reported in this chapter was partially developed by researchers involved in the original study. An independent team of researchers would be desirable in order to avoid any bias from the potential conflict of interest researchers could hold with regard to seeing a successful replication of their own study. But in this case, the data was collected for the analysis presented in chapter 3 initially was not intended to be used for a direct replication. The opportunity to attempt a replication with such closely matched methods represented an excellent opportunity to contribute to the robustness of literature in the field, in the light of the debated issues with replicability (Gilbert, King, Pettigrew, & Wilson, 2016).

Although many of the important aspects of this replication are matched to the original study, an aspect where they both differ is in the time frame of data collection. The original study selected all relevant press releases for the year

2011, but since the data for this replication were originally intended to be used in an interrupted time-series design, the data was collected for the period of January to June in 2014 and 2015. This means that any variance in press releases and news stories in the second half of the year would not have been included in the replication. It is not clear whether there would be any difference in the selection of stories by the newsrooms, or whether news content would change from July onwards, but the output of press offices at the end of the year may be slower than at the start of the year. In the replication sample, there were 522 relevant press releases (before being restricted to a maximum of 20 per institution), in comparison to 462 relevant press releases in the original study. This increase could be just an increase in output year-on-year, or it could be indicative of a differential within the year. Subjects such as psychology (which often bases research on samples of term-time students) for which health-related data might not be collected all year around may have contributed unequally to the replication dataset than to the original study. If indeed there is such a differential, matching the original study's method of collecting data over the course of a single year would be a safer approach for future replication attempts.

2.5.3. Conclusion

The findings of this replication provide support for the main conclusions from Sumner *et al.* (2014). Press release characteristics appear to be largely similar to the previous research, as does the relationship between the press release and the news article. Though exaggeration can originate in news articles, it is more likely to be present if the same type of exaggeration occurs in the press release. These findings are purely observational, and need to be supported by experimental

evidence in order to be able to say that exaggeration earlier in the science communication chain directly leads to exaggeration in news articles.

Nonetheless, these findings should be seen as impetus for scientists and press officers to follow the press release labeling guidelines set out by the Academy of Medical Sciences (2018), and “to communicate research accurately, without over-stating results and misleading the public – particularly when it comes to health” (Stempra, 2017), there appears to be no benefit to exaggerating results in terms of media coverage – only the danger that the public may be misinformed.

Differences in the levels of advice seen in news, and potentials in press releases, suggests that there may be a growing trend for the inclusion of advice, potentially due to the continuing expansion of Internet access. The more cautious consideration of the inclusion of advice should become a priority of the science communication community to try to mimic the potential success of the Concordat on Openness on Animal Research in the UK (2014).

This study differs to that of Schat, Bossema, Numans, Smeets, and Burger (2018) in that Dutch news uptake was higher for articles that contained exaggerated statements of relationship, whereas there was no difference in this study. It is unclear why Dutch newsrooms would be more likely to pick stories that are exaggerated. The authors themselves only suggested that the difference in uptake between exaggerated and non-exaggerated stories was small, but that it may be useful to examine whether there is a difference in English and Dutch newsroom selection criteria. It seems unlikely that there would be a notable difference in the journalistic practices between the UK and the Netherlands in general. A point of

interest is that the majority of press releases used by Schat, Bossema, Numans, Smeets, and Burger (2018) were from university medical centers, with a minority of press releases being sourced from general university press offices. There appears to be a trend for health-related research to be conducted and communicated by such specialised institutions. So the question would be whether this factor modulates the type of research output, and whether this type of research is communicated differently. Schat, Bossema, Numans, Smeets, and Burger (2018) did not provide report the difference in exaggeration across press office type, but this would be a useful area of future enquiry.

CHAPTER THREE – INTERRUPTED TIME SERIES

3.1. Introduction

Chapter two described the attempt at a replication of the main findings of Sumner *et al.* (2014) that there was a relationship between the level of exaggeration in press releases, and the level of the same exaggerations in related news articles. The replication followed the original study as closely as possible and found that the findings are largely consistent in a dataset collected a few years later. This chapter is concerned with using that same dataset, but with updated methods, to see whether there was any change in the level of exaggeration in journal articles, press releases, and news articles after the publication of the paper in The BMJ in later 2014.

3.1.1. Paper impact

Since its release in December 2014, the Sumner *et al.* (2014) article has been linked to around 40 news articles, 50 blog posts, and tweets from around 1500 users and is ranked in the top 5% of all research outputs by Altmetric (Altmetric, 2018), as well as being cited in around 150 academic articles. Given this wide reception, it is possible that the article's implications and recommendations for press officers might have had an impact on behaviour. In the editorial for the paper, published in the BMJ, Ben Goldacre called for accountability in academic press releases, to remedy misrepresentations in scientific findings, as well as calling for transparency through which press releases would be more strongly aligned to the infrastructure surrounding original research article, in full view of peers, rather than being sent privately to news organisations (Goldacre, 2014).

Given the pay-walls on many journals, people interested in new health information have to rely on press releases, or information from press releases filtered through news (Young, 2017).

Given that press offices could hold a high level of influence on news content and public understanding, the findings of Sumner *et al.* (2014) provided implications for the practice of press offices. Guidelines published by the Academy of Medical Sciences (2017), citing Sumner *et al.* (2014) recommended that press officers should “be clear about whether the reported finding is a correlation or causation”, by using causal language only when the research can support such conclusions with a suitably strong methodology. Press officers were some of the most important stakeholders of the findings of Sumner *et al.* (2014), and the spotlight fell on them, not entirely, but to a large degree, to consider their practices.

3.1.2. Causal language

The conflation of correlation and causation is a perennial topic in academia and science reporting. Norris, Philips and Korpan (2003) showed that undergraduate students understood statements of relationship to be stronger than they actually were, despite it being a fundamental skill in science to distinguish between correlation and causation. Examining data from the dataset used in the previous chapter, around 32% of all news articles reporting observational research contained explicit statements of cause. With such frequency, it is relatively common to see such spurious statements of causation in the media. For example, in 2012, an article by the Daily Mail stated “Violent video games 'make

teenagers more aggressive towards other people”, before also stating a few lines later “surveys were carried out across four school years”. The study was indeed based on self-report surveys (Willoughby, Adachi, & Good, 2012). As an observational research method, such surveys cannot provide the type of evidence that would justify a causal statement between variables. In the sentence “Violent video games 'make teenagers more aggressive towards other people’”, the variables ‘violent video game use’ and ‘aggressive behaviour’, are linked by the causal phrase ‘make teenagers more’, suggesting that video game use causes the aggressive behaviour. This type of exaggeration of statements of relationship is relatively common.

Some of the most high profile cases of scientific inaccuracies are due such conflation of correlation and causation. In perhaps the most famous example of such inaccuracy, the discredited and retracted report of a causal link between vaccines and autism still persists as a popular stance (Poland & Spier, 2010), which could result in devastating consequences, such as the reduction in rates of vaccination, and the spread of virus (Sugerman *et al.* 2010). This is of course based on a self-report observational study (with a low sample size and selection bias, by an author with a serious conflict of interest) that has since been retracted and has been found to have used fabricated data (Godlee, Smith, & Marcovich, 2011). The claim that vaccines cause autism is of course a causal claim, and an exaggeration of the evidence. It is unlikely that the same detrimental outcomes to public understanding and public health would have been recorded if the headlines read ‘vaccinations associated with onset of Autism’, or more representative of the current evidence: ‘no link between vaccines and autism’.

This is of course and extreme example. Causal relationships are frequently reported in the news, but it is very uncommon to witness a false health scare such as this.

A strict adherence to the rule of using associative statements for relationships uncovered by observational methods, and causal statements for relationships discovered using experimental methods could help to avoid such misrepresentations. That is not to say that all observational studies are too weak to support an inference of a causal claim. Many researchers would probably agree that the finding that smokers had a higher incidence of mortality due to lung cancer, and that this was a dose-dependent effect, in a sample of 40,000 participants (Doll & Hill, 1956), was a strong finding which could justify a causal claim. Rubin (2007) suggests that in certain fields that are not conducive to research with experimental methods (such as investigation of substance abuse in humans), strong observational procedures can be designed to approximate their causally inferring counterparts. That being said, generally observational studies would not meet these standards, and classifying statements of relationship by study design may be useful.

3.1.3. Categorising statements of relationship

The InSciOut research group at Cardiff University produced the categories of statements of relationship used to calculate exaggeration in Sumner *et al.* (2014) by a consensus method. The research team generated a list of statements and categorised them in order of strength until seven categories emerged (table 3.1). Ranked from 0 to 6 of increasing strength, the system included correlational

statements two (such as ‘associated with’, and ‘has higher rates’) at rank 2, and causal statements (such as ‘influenced’, and ‘prevented’) at rank six.

Correlational statements were classified as those that can be reversed and their meaning remains the same: for example ‘x is related to y’. Between the correlational and causal categories were statements classified as ambiguous, which can imply more information than a correlational statement (‘linked to’, ‘predicts’); conditional cause statements (such as ‘might cause’, or ‘could cause’), which contain a modal verb than acts as a lexical hedge giving uncertainty to the causal statement that precedes it (Hyland, 1996); and ‘can cause’ statements, which express certainty that causation occurs under certain circumstances. Sumner *et al.* (2014) classed exaggeration as the inclusion of a statement in any higher category than the statement in the source article – so ‘elevated testosterone could increase fatigue’ (a conditional cause statement) in a press release would be an exaggeration of ‘fatigue increases with testosterone levels’ (a correlational statement) in a journal article.

Further research by others in my research group demonstrated that these literal interpretations of the strength of relationship, and the readers’ perceived strength of relationships are slightly different. The order of strength of relationship of statements introduced by Sumner *et al.* (2014) was preserved, but some of the categories of classification were not perceived to be different (Adams *et al.* 2017). Adams *et al.* (2017) tested participants’ comprehension of such statements of relationship by asking them to rank a series of statements in order of the degree of causal implication. This method revealed that not all categories were perceived to be distinct from each other. Correlational, ambiguous, and

conditional cause statements were not perceived to be distinct categories by readers, but ‘can cause’ was perceived to have a higher degree of causal implication, and ‘cause’ to have a yet further degree of causal implication. This finding categorised ambiguous statements and conditional cause statements to be of equal strength to correlational statements (demonstrated in the grey box in table 3.1). This left five categories, in descending strength order: 5) causal statements; 4) ‘can cause’; 3) ‘conditional cause’, ambiguous statements, and correlation; 2) does not cause; and, 1) no cause mentioned. In a reanalysis of the findings of Sumner *et al.* (2014) and Sumner *et al.* (2016), Adams *et al.* (2017) demonstrated that this re-categorisation of statements of relationship yielded lower rates of exaggeration. Given that this categorisation is based on reader perception, it provides a more valid interpretation of exaggeration, because the rates of exaggeration calculated are likely to be closer approximations of the exaggeration perceived by readers of news articles.

Table 3.1.

Coding categories for statements of relationship showing the interpretation by Sumner et al., (2014), and the later modification by Adams et al., (2017). Categories are listed with descending strength, one category per line. The shaded area shows the categories that were found to be equivalent in the readers’ interpretation by Adams et al., (2017).

Sumner <i>et al.</i> , 2014	Adams <i>et al.</i> , 2017
Cause	Cause
Can cause	Can cause
Conditional cause	Conditional cause
Ambiguous	Ambiguous
Associative	Associative
Does not cause	Does not cause
No cause mentioned	No cause mentioned

3.1.4. Present study

In order to make comparisons of the levels of exaggeration in journal articles, press releases, and news, before versus after the release of Sumner *et al.* (2014), the dataset used in the replication in chapter 2 was used. This dataset was originally planned to allow for this dual purpose. Since Sumner *et al.* (2014) used data from the 20 universities that were members of the Russell Group in 2011, it would be expected that these universities would, as the major stakeholders of the findings of Sumner *et al.* (2014), be most likely to change their behaviour in line with the study's findings. For this same reason, a sample of press releases published by The BMJ was included in the analysis. The BMJ is a prominent medical journal, and it publishes a high volume of health-related research, as well as associated press releases. The BMJ published the Sumner *et al.* (2014) paper. Given the implications for practice published in that article, and the subsequent magnitude response, as indicated by Altmetric, and as discussed in the editorial by Ben Goldacre (2014), it is expected that the content of press releases published by The BMJ would also change.

In the following analysis, journal articles, press releases, and news articles were scrutinised to see if statements of relationship were used more accurately after the publication of Sumner *et al.* (2014). In other words, was there a reduction in exaggeration in articles in 2015 versus 2014? It was expected that press offices, given the focus on them in the implications and follow up to the Sumner *et al.* (2014) article, would contain less exaggeration in the 2015 data, than the 2014 data. For news articles and journal articles, it was unclear whether the implications of the research would have had a direct impact. For journal articles,

there may be some reduction in exaggeration due to the relationship between press offices and scientists. For news, it was expected that there would be an overall reduction in exaggeration, given a reduction in the proportion of exaggerated press releases. This analysis also explored whether exaggeration in news reduced independently of the content of press releases.

3.2. Method

This study shares much of its data and design with the replication analysis described in chapter two. Additional information regarding the nuances of the data collection and coding process is provided below. Some details were omitted from chapter 2 because they were not relevant to that design. The two studies diverge in their interpretation of exaggeration, and in their analyses. The previous study used the Sumner *et al.* (2014) interpretation of exaggeration, whereas the present study used the updated interpretation of Adams *et al.* (2017) (table 3.1). The analysis was performed based on cases where observational journal articles made correlational or equivalent statements of relationship (explained below).

3.2.1. Data collection

The same data reported in the replication analysis in chapter two were used here, in addition to data from press releases published by The BMJ. The sample consists of 371 press releases: 351 from the 20 universities that were members of the Russell Group in 2011 (the sample of Sumner *et al.*, 2014), and 20 from The BMJ.

Selection was performed as described in chapter two. Press releases for the period of January to June 2014 and January to June 2015 were collected from the press offices' web pages, and from EurekAlert.org. There were 230 press releases published by The BMJ across January and June 2014, and January to June 2015. Of these, 143 were relevant in that they reported on published, peer-reviewed research, with relevance to human health. This sample was restricted to 10 press releases from January to June 2014, and 10 press releases from January to June 2015 via a randomisation process. This restriction is to control the extent to which data from a single press office contributed towards the dataset, and to limit the time the data set took to create via the consensus coding method employed. Figure 3.1 demonstrates the selection process for the sample. This is the same as figure 2.1, but with the addition of press releases from The BMJ.

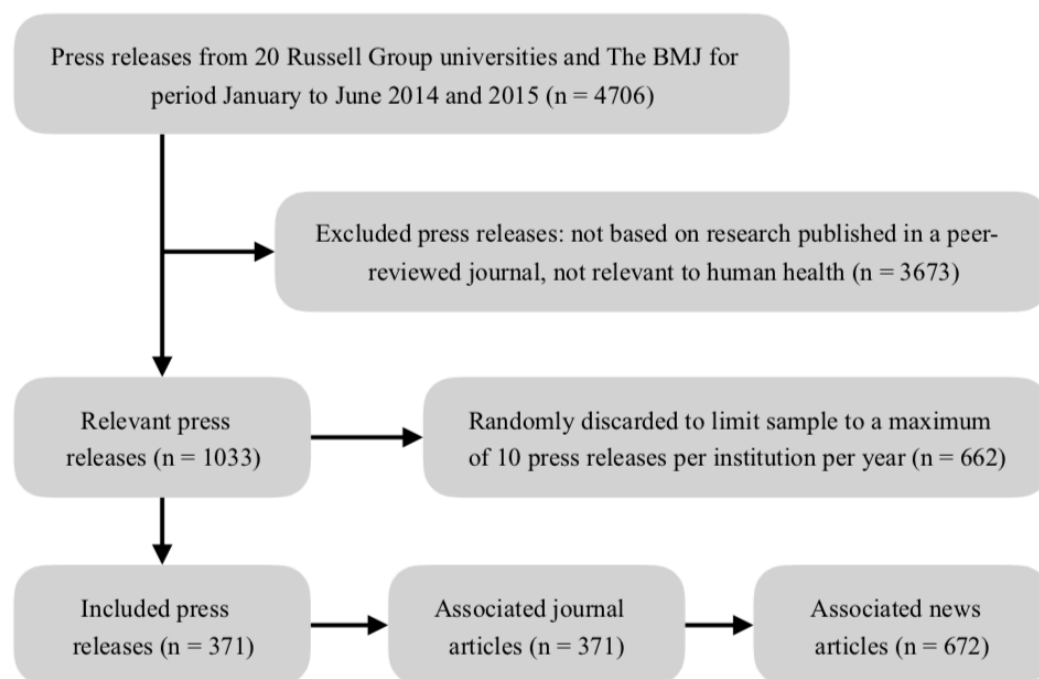


Figure 3.1. Press release and news selection diagram.

For each press release, associated news articles were collected from Google search, and the Nexis database (LexisNexis, New York, NY) using keywords, such as the variables used in the study. This search was conducted up to 28 days after publication and up to a week before to account for any news article being released before the embargo was lifted on the press release.

3.2.2. Article coding

Prior to the coding method outlined in chapter 2, the corpus of articles underwent a redaction process to remove any references to the year 2014 or 2015. This was so that the coders, who were aware of the aim of the study, were not aware which condition the articles they were coding belonged to. This redaction process was achieved using Automator software (Apple Inc.). The articles were searched for information of interest, and coded using the sheet provided in appendix 2.1. For this analysis, only information regarding the statements of relationship reported in each article was used. For each article set, two researchers independently coded each article, and subsequently both researchers' work was compared electronically for disagreements. The coders then met to remedy disagreements flagged by the comparison of their articles, with the help of a third coder if a disagreement was difficult to resolve. This created a database with 100% agreement in coding. No data was collected regarding the proportion of agreement between coders before consensus, but the previous research found between a 91% and 98% agreement between coders. The consensus method used here eliminated disagreement completely.

3.2.3. Analysis

Exaggeration was defined using the category structure reported by Adams *et al.*, (2017). There were five categories of increasing strength ranging from 1) no cause mentioned, 2) statement of no effect, 3) correlational or equivalent (including ambiguous, and conditional cause such as ‘may’ cause, ‘might’ cause, or ‘could’ cause), 4) can cause, and 5) cause. Categories 3 (correlational) and lower are interpreted to be appropriate statements when referring to the findings of observational designs, with categories 4 (can cause) and 5 (cause) being appropriate for experimental designs. When calculating exaggeration between two statements, a statement in category 4 or 5 would be deemed an exaggeration in comparison to a statement in a category below. There was no differentiation made between statements 4 and 5.

Generalised estimating equations (GEE) were used to generate percentage rates of exaggeration and 95% confidence intervals, with exchangeable working correlation to adjust for clustering of multiple articles to one source. The binary distribution was specified, as all outcomes were binary, with a logit linking function to enable easy interpretation of odds ratios.

For the calculation of journal exaggeration, the strongest statement of relationship in the title, abstract, and conclusions was compared to the study design. The sample was limited to cases where the study design was observational cross-sectional, observational longitudinal, or an observational meta-analysis. Rates of exaggeration were calculated as cases where the journal reported a relationship between variables in category 4 (can cause), or 5 (cause),

as these are not justified by the study design. The year of the sample (2014 or 2015) was entered as a predictor into the model, and the press office institution was specified as the subject variable upon which journal article exaggeration is clustered. There were 168 cases available for analysis.

For the calculation of press release exaggeration between 2014 and 2015, the strongest statement of relationship present was compared to the journal study design. Statements of relationship were considered if they appeared in the title or first two sentences of the body of the article that were not context. The year of the sample was entered as a predictor, and press office was specified as the subject variable.

For analysis of exaggeration in news articles, the strongest statement of relationship from the title, or main statements of the body of the news article, was compared to the journal article design. Only cases where the journal article did not exaggerate were included in the analysis. That is, cases where the strongest statement in the journal article was no stronger than a correlational statement, for observational study designs. This means that exaggeration in news articles could be compared to study design without including cases where journalists could have been exposed to exaggeration originating in journal article statements. Whereas press officers should have a close relationship with the scientists involved in the study being reported, journalists are much less likely to have such contact. It would be expected that for this reason, and because the press office has been implicated in the origin of exaggeration by previous research, that press releases should be aligned to the study design. News

exaggeration as recorded using this interpretation will originate in the press release, or the news itself. There were 322 cases available for analysis of news exaggeration. The journal article was specified as a subject variable in order to account for clustering of multiple news articles reporting on a single story, and the year of the sample was entered as a predictor.

For the analysis of the relationship between exaggeration in press releases and news articles, exaggeration in the news article was again compared to non-exaggerated observational journal articles. The journal article was specified as the subject variable to account for clustering of multiple news articles to one story, and the interaction between press release exaggeration and the year of sample was entered as a predictor. This analysis will demonstrate the news reliance on press release content between 2014 and 2015.

3.3. Results

3.3.1. Journal article exaggeration

For the comparison of exaggeration in statements of relationship in journal articles, in comparison to the journal design, between 2014 and 2015, 40% (95% confidence interval = 35% to 45%) of journal articles made a claim that was in excess of that justified by the observational study design. The odds of such exaggeration were not found to be higher in 2014 (45% of articles exaggerated, 95% confidence interval = 36% to 55%) than 2015 (34% of articles exaggerated, 95% confidence interval = 27% to 42%): difference = 11%, 95% confidence interval = -1.6% to 23.6%; odds ratio = 1.6, 95% confidence interval = .9 to 3.0.

3.3.2. Press release exaggeration

For the comparison of exaggeration in press releases across 2014 and 2015, 37% (95% confidence interval = 29% to 45%) of press releases made a claim that was in excess of that justified by the study design of the original research. The odds of such exaggeration were higher in 2014 (45% of articles exaggerated, 95% confidence interval = 35% to 56%) than 2015 (29% of articles exaggerated, 95% confidence interval = 21% to 38%): odds ratio = 2.1, 95% confidence interval = 1.2 to 3.5; difference between conditions = 16%, 95% confidence interval = 2.1% to 29.9%. Because the confidence interval of the difference between exaggeration in 2014 and 2015 does not include 0, this is indicative of exaggeration being predicted by year. Thus, exaggeration was lower in 2015.

3.3.3. News exaggeration

For the comparison of exaggeration in news across 2014 and 2015, 39% (95% confidence interval = 28% to 52%) of news contained exaggeration. The odds of such exaggeration in news were no higher in 2014 (41% of articles exaggerated, 95% confidence interval = 27% to 57%) than in 2015 (37% of articles exaggerated, 95% confidence interval = 23% to 55%): odds ratio = .9, 95% confidence interval = .3 to 2.1; difference between conditions = 4%, 95% confidence interval = -18.5% to 26.5%. Figure 3.2 demonstrates levels of exaggeration for each journal articles, press releases, and news articles, between 2014 and 2015.

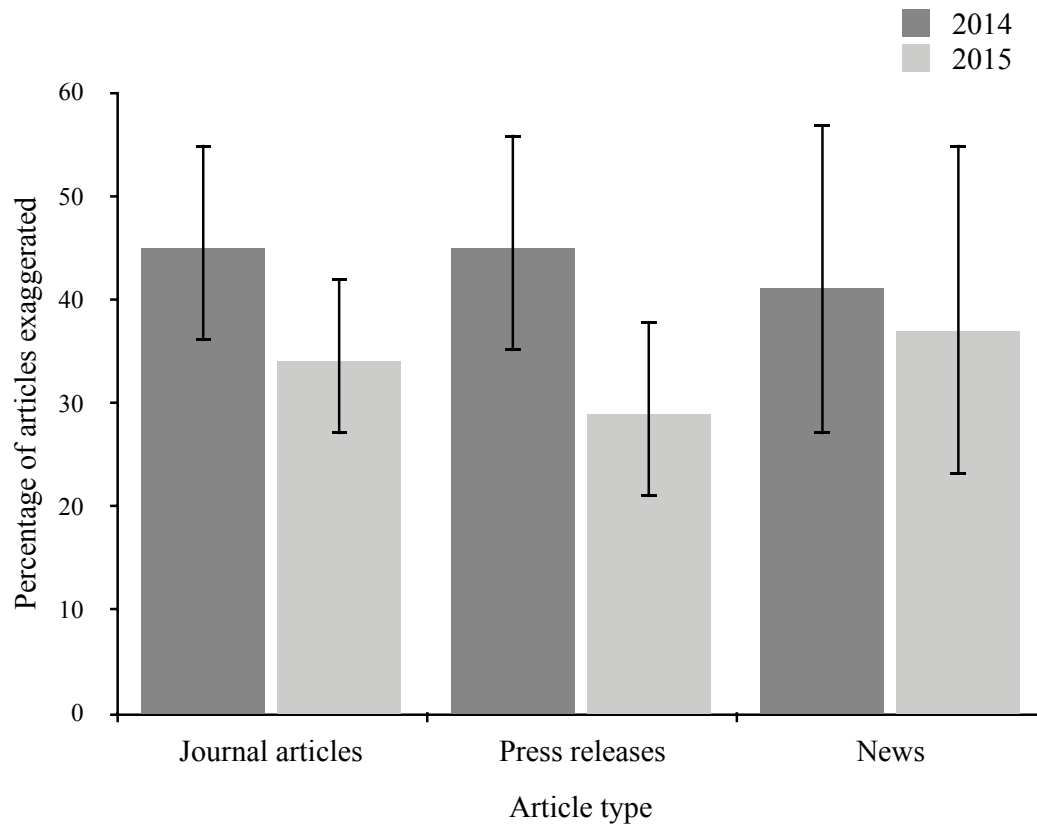


Figure 3.2. Exaggeration in journal articles, press releases and news between 2014 and 2015. Exaggeration in press releases was significantly reduced 2015 versus 2014. Error bars represent 95% confidence intervals.

3.3.4. News exaggeration as a function of press release exaggeration

The odds of exaggerated news were 10.5 times higher (95% confidence interval = 3.0 to 36.8) for exaggerated press releases (70% exaggerated, 95% confidence interval = 49% to 85%), than for non-exaggerated press releases (16% exaggerated, 95% confidence interval = 10% to 24%) across 2014 and 2015 combined; difference between conditions = 54%, 95% confidence interval = 34.0 to 74.0. For exaggerated press releases, the odds of exaggeration in news were no higher in 2014 (69% of articles exaggerated, 95% confidence interval = 44% to 86%), than in 2015 (70% of articles exaggerated, 95% confidence interval = 33% to 91%): odds ratio = 1.1, 95% confidence interval = .2 to 6.2; difference

between conditions = 1%, 95% confidence interval = -36.3% to 38.3%. Figure 3.3 demonstrates the levels of exaggeration in news between 2014 and 2015, compared between press release exaggeration.

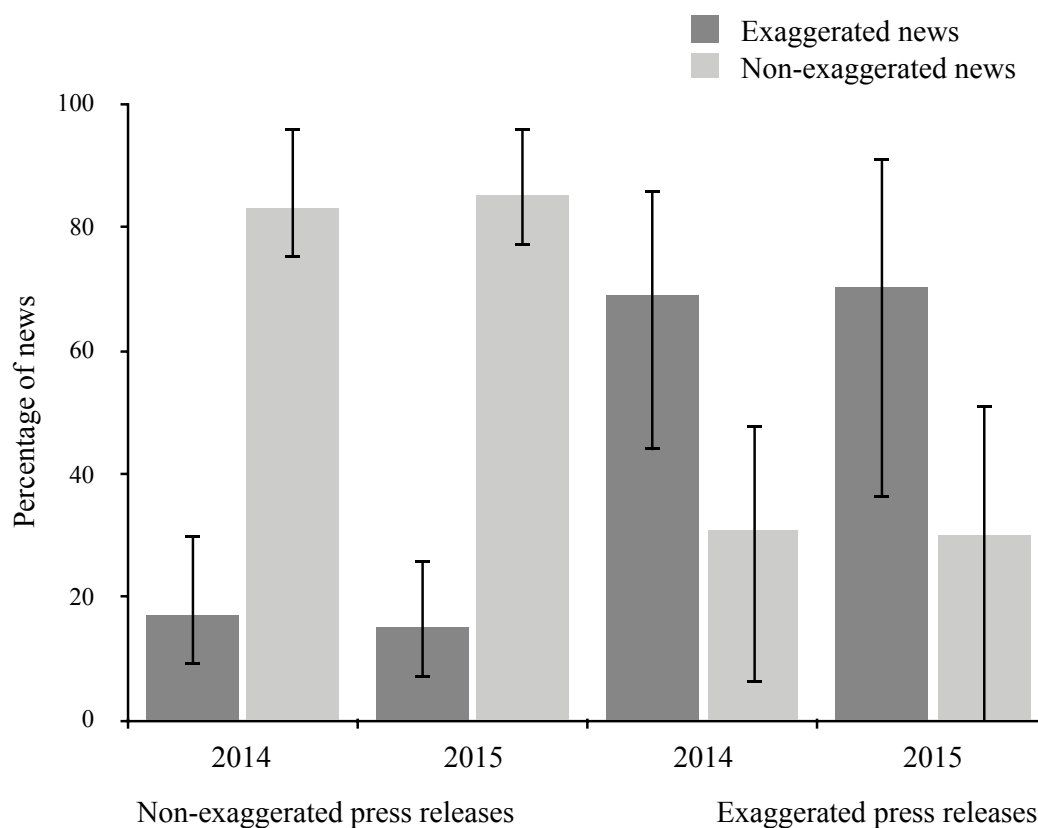


Figure 3.3. Exaggeration in news articles as a function of press release exaggeration and year of publication.

3.4. Discussion

The aim of this study was to examine whether exaggeration in the reporting of observational health-related research findings had reduced after the release of an influential paper in the area. Levels of exaggeration in news and journal articles were not found to be significantly lower in 2015 than in 2014, but the overall level of exaggeration within observational research papers (35% in 2015) is

unexpectedly, and worryingly high. The level of exaggeration in press releases was lower in 2015, suggesting that press officers had become more cautious when reporting the results of observational research. Although news exaggeration was predicted by press release exaggeration over the whole sample period (confirming the findings of chapter two with amended methods) and this relationship was similar in both years, the overall exaggeration rates in news did not change enough in 2015 versus 2014 to be significantly detectable.

Given that the findings of Sumner *et al.* (2014) had the biggest implications for press offices, the detected change in their output and the lack of a detected change in journal articles and news could be explained in a number of ways. Firstly, journalists and scientists are one step removed from the research in terms of its implications compared to press officers. Even though Sumner *et al.* (2014) were cautious with their conclusions in not apportioning blame to any single party, scientists and journalists are implicated by the findings to a lesser degree than the press officers. Whereas the blame for exaggerated reporting could be easily aimed at journalists prior to the paper, the magnitude of the relationship between press release and news exaggeration turned the focus to press offices.

Scientists are still involved in the process of creation of press releases, given that they contribute and have a final say before release, but it is unlikely that their contribution changed between 2014 and 2015 if exaggeration in their own papers did not change. Journalists are known to be more likely to print exaggeration when the press release contains exaggeration, and this finding was echoed here, but the lack of a detected change to exaggeration in news despite the change in

press releases is either indicative of a failure of the research to detect a change, a lack of an actual causal link between press release exaggeration and news, or the unlikely occurrence of another parallel event which had an equal and opposite effect to the release of the Sumner *et al.* (2014) paper. A possibility is that the further communication that occurs between actors in the science news process may be responsible. Even though the press releases were found to contain less exaggeration, this study could not record the interaction between journalists and press officers, such as other information in the emails sent from press officers to journalists with press releases attached, and any further conversation. Journalists are likely to either reach out to the press office, the researchers, or other experts or sources to gain more information. These extra interactions, which are not as prescribed as the press release, and that are probably more spontaneous, will be likely to contain information that is less considered and accurate than that found in the press release. It could be that a greater proportion of the information present in news stories is actually based on further interaction, and independent investigation, as opposed to churnalism.

Actors at all stages of the science media process can still improve their practice, but to a lesser degree than press officers. Press officers may have changed behaviour because they are in the spotlight. Some scientists and journalists are essentially one step removed from the focus of the problem, so they lack the same impetus to change. In addition to the focus of the findings being on them, another reason for the change in behaviour of press officers may be that during the January to June time-frame of 2015 (the second half of the sample in this study) many press officers were aware of, or had been directly contacted

regarding, the randomised controlled trial in which the InSciOut team were planning to modify and monitor the effect of press office outputs on the news. Knowledge of the potential to be observed may have manifested as a pre-Hawthorne Effect (McCarney, *et al.*, 2007)

3.4.2. Limitations

The retrospective and observational nature of this study means that it would be a conjecture to state any change in behaviour is due to the publication of Sumner *et al.* (2014). We cannot be sure whether other events during the interruption in the sample may have contributed to or caused a change in behaviour. Such observational research is only suggestive of a link between exaggeration transfer from press releases to news, as discussed in the last chapter, and is only suggestive of an influenced change in behaviour. The consciousness regarding the misleading content of science reports had been developing over a number of years with such findings reported in a number of other studies, such as Schwartz, Woloshin, Andrews, & Stuckel, (2012); and Brechman, Lee, and Cappella (2009). There are a few potential routes for support of the present findings. One approach would be to code the same articles used in this sample for a type of exaggeration, or quality measure that was not highlighted by Sumner *et al.* (2014), such as the presence of information about absolute risk, as studied by Schwartz, Woloshin, Andrews, and Stuckel, (2012). Testing for changes in related measures that were not reported in the quasi-intervention used in the present study could reveal whether the change detected is an overall improvement (likely due to other factors), or just an improvement in exaggeration of statements of relationship (more likely related to the quasi-

intervention). Another route for confirmation would be to examine changes in behaviour with experimental methods. This is the approach reported in the rest of this thesis.

3.4.3. Conclusion

The conflation of correlation and causation is still evident, and common in science reports. Regardless of whether the release of the paper by Sumner *et al.* (2014), or any parallel occurrences led to the reduction in exaggeration of statements of relationship in press releases, the finding should be taken as another positive example of improvement in the reporting of science, just as with the improvements to reporting of animal research. If the findings are indicative of an effect of the research, this is a positive tale of impact.

The present research did not reveal a change over time in the relationship between press releases and news, where exaggeration in news is elevated following exaggerated press releases. Press offices should be encouraged to continue to improve the accuracy of their output given that there appears to be no penalty to newsworthiness. Given that this was observational research, any interaction between the actors in the science communication process outside of the published articles analysed cannot be controlled. The experimental research outlined in chapters five and six control for such extraneous variables so only the potential influence of the source articles will be detected. Outside of the present thesis, future research could investigate the interactions between scientists, press officers, and journalists. It is difficult to directly test participants in those job roles, but there is potential for a survey to collect information regarding how

frequently each actor interacts with another in the science communication process, and to what extent they obtain information that ends up in final articles from these interactions.

As a future direction for further research, the worryingly high level of exaggeration within observational journal articles should be scrutinised. Research articles are often inspected by peers, before being formally reviewed by journals before being published, so it would be assumed that either the exaggeration present seemed suitable, or more likely that it was not detected, or not deemed to be exaggeration. An interesting approach to this might be to create a corpus of the research articles that have been found to contain exaggeration and to survey the article authors to determine 1) whether the authors believe that exaggeration exists within their article; and, 2) whether the authors agree with the definitions of exaggeration used in this study, or test their comprehension of the levels of statements of relationship using the method of Adams *et al.* (2017). This data could then be compared to that of authors of non-exaggerated articles, and this should allow for inferences to be made about whether authors perceive exaggeration as defined by this research, and whether they detect exaggeration.

CHAPTER 4 – RANDOMISED CONTROLLED TRIAL

4.1. A note on contribution

This chapter reports on a randomised controlled carried out by the InSciOut research group. The research was initially formulated and funding was secured prior to my addition to the research group. I was involved in the project from the data collection phase onwards. Primarily I was jointly responsible for data coding and data handling prior to analysis. I also supported the day-to-day running of the trial protocol, and assisted with analysis and interpretation. The following is my account of the research.

4.2. Introduction

Randomised controlled trials are the gold standard for making causal inferences. Whereas a rigorous and well-constructed observational study, with many relevant recorded covariates, can lead to correct inferences (Rubin, 2007), the randomised controlled trial is widely accepted as the best way to infer cause and effect, given the properties of randomisation to conditions, blinding, and adequate control (Sibbald & Roland, 1998).

The previous work on examining exaggeration in science reporting has relied heavily on observational methods. Typically articles from different stages of the science media process are assessed to calculate the proportion that contain certain exaggerations. The research reported here aimed to intervene in the process via press releases (the participants in the trial), to measure outcomes in

the content of news stories. It was hoped that this approach would strengthen the inferences made in previous research (Sumner *et al.*, 2014; Sumner *et al.*, 2016).

In addition to examining whether changes to the strength of statements of relationship would be reflected in the news, the study also aimed to measure whether important caveats regarding the research design made it through to the news. Caveats to research design are statements that qualify the claims made with regard to the strength of the underlying methodology. For example, for observational research, a suitable caveat would be ‘given the observational nature of this research, we cannot infer cause and effect. Further experiments would need to be conducted to show whether chocolate consumption directly reduces blood pressure’. Such caveats have been shown to be present in the news when they are present in press releases, but their presence in press releases is rare (Sumner *et al.*, 2014).

As shown in chapter two, there is uncertainty regarding whether exaggerated press releases were more likely to be reported in the news. There was no difference in news uptake detected by Sumner *et al.* (2014), and my replication with more recent data, though a difference has been found elsewhere (Schat, Bossema, Numans, Smeets, & Burger, 2018). This trial aimed to see whether an intervention to press releases, changing statements of relationship to be more aligned to the research design, would have an effect. No reduction in news uptake would be seen as no penalty for improving the alignment of statements to those justified by the study design. Similarly, would the inclusion of caveats alter the likelihood of articles being reported in the news? Caveats often highlight

limitations of the research, and could reduce the readers' confidence in the story, but might not affect interest, or the likelihood of the article being selected for news (Bott *et al.*, 2018).

4.3. Method

4.3.1. Participants

Between September 2016 and May 2017, nine press offices in the UK published 312 press releases that had undergone intervention as part of this study. Press releases reported on biomedical or health-related research findings published in peer reviewed research articles – the press releases acted as participants in the trial. The press offices were invited to send any press release they perceived to be eligible; during the trial they generated 622 press releases in total. Table 4.1 demonstrates the reasons for exclusions at the level of the press offices, and the research team. The criteria for eligibility were: consent from the research author, relevant to human health, and that the press office led the press release (in the case of joint press releases). Then, only press releases reporting on observational (cross sectional or longitudinal, and observational meta analysis) and experimental (randomised controlled trial, experimental meta analysis, and other experimental) research were accepted for random assignment to conditions. A total of 2257 related print, online, and broadcast news articles were collected via searches on Google, Nexis (LexisNexis, New York, NY), and TVEyes (TVEyes, Fairfield, CT).

Table 4.1.

The numbers of available press releases excluded from the trial, with reasons for exclusion.

	Number	Reason
Total press releases available	622	-
Excluded by press office		-
	-100	Joint press release - not lead press office
	-52	Overlooked or not specified
	-48	Time constraint
	-23	Author did not consent
	-21	Ineligible topic
	-17	Staff turnover/absence
Press releases submitted to trial	361	-
Excluded by research team		-
	-48	Not relevant study design
	-1	Missed by researchers
Remaining for intervention	312	-

4.3.2. Design and procedure

A 2x2 design was employed where press releases were randomly assigned to either the control manipulation, or one of three experimental manipulations: change to causal claim, change to caveat, or change to both (table 4.2). The factors were the presence or absence of caveats, and the presence or absence of a manipulation to the causal claim. In the condition containing no inserted caveat, and no change to causal claim, a synonym change was added to a suitable word in the first paragraph. Suitable words were deemed to be those that would not be altered in any other manipulation. This word was added as a placebo manipulation, to ensure that participants in the control condition still underwent the same process of submitting and reviewing a draft press release prior to final release; in other words, so that all conditions were treated the same apart from the experimental manipulations. In the causal claim condition, headlines and main statements were altered to better reflect the underlying research design.

This was done in accordance with the categories of statements of relationship indicated by Adams *et al.* (2017), as used in chapter 3. Correlational or equivalent statements were inserted for press releases with observational designs, and causal or ‘can cause’ statements were included for experimental studies. For caveats, an appropriate statement was inserted that conveyed the strength or limitation of the study design with regard to the ability to infer cause and effect. For press releases assigned to the control condition a synonym was inserted for a word present that was not related to the statement of relationship or caveat. The final condition contained both modifications to statements of relationship, and caveat insertion. If any press release assigned to an experimental condition already contained the proposed alterations, it was unchanged and entered into the analysis in its original state. Press offices were always given the final say in whether changes would be accepted.

Table 4.2.

Conditions for trial interventions showing the number of press releases randomly assigned to each category.

Condition (alteration made)	N
Control/synonym change	89
Statement of relationship change	64
Addition of caveat	79
Statement of relationship change and addition of caveat	80
Total N	312

4.3.3. Article coding

Press releases, source journal articles, and subsequent news articles were coded using the same protocol to that outlined in the previous two chapters. The information of interest was the strength of the statements of relationship in the titles and the first two main statements of the articles, and the presence of

caveats. Coders were blind to assignment of the articles during the coding process.

4.3.4. Analysis

A relatively high proportion of press releases submitted to the trial already contained the changes of interest. For this reason, analysing data only from press releases that were published with the condition-assigned changes intact (per-protocol analysis) was not possible. Two types of analysis were therefore performed: an as-treated analysis (AT), in other words, an observational analysis of the actual press release content regardless of condition; and an intention-to-treat analysis (ITT), where all press releases were treated as if they had adhered to the assigned condition manipulations. Generalised estimating equations were used as in the previous chapters to obtain confidence intervals for outcome variables whilst adjusting for multiple news articles to one press release.

Outcome measures were the proportion of articles that contained aligned claims and the percentage of press releases that had related news. Aligned claims are those that contained a statement of relationship that was suitable for the study design. In other words, articles reporting on observational studies that contained correlational or equivalent language; or articles reporting on experimental research that contained causal, or ‘can cause’ statements. Headlines and main statements of news articles were analysed separately because they are usually written by different people.

4.4. Results

4.4.1. Intention-to-treat analysis – statements of relationship

For this analysis, press releases that were assigned to the synonym change group (control), and the caveat group were treated as controls (because they contained no alterations to statements of relationship), and the two conditions that contained alterations to statements of relationship were the experimental groups. There were 168 press releases in the control condition, and 144 press releases in the experimental conditions. Of those in the experimental condition, 24 press releases did not contain the suggested change, 30 had no available relationship for manipulation, and 55 had already included an accurate statement. This meant that only 35 out of 144 press releases contained the correct experimental manipulation to causal statements. This ITT analysis, including all cases regardless of condition adherence, demonstrated no significant difference in the number of headlines that were aligned to the research design (95% confidence interval of the odds ratio = .7 to 1.6), and no significant difference in the number of aligned main statements (95% confidence interval of the odds ratio = .7 to 1.6).

For news uptake, the same analysis was performed but with a binary measure of news uptake as the outcome variable, rather than news alignment. There was no difference in news uptake for stories reported in press releases in the conditions that contained alterations to statements of relationship versus the control conditions: 95% confidence interval of the odds ratio = .7 to 1.3

4.4.2. Intention-to-treat analysis – caveats

For this analysis, press releases that were assigned to the synonym change group (control), and the statement of relationship change group were treated as controls, and the two conditions that were assigned the addition of caveats were the experimental groups. There were 153 control press releases, and 159 experimental press releases. Of these 159 press releases, 45 did not contain the suggested caveat, and 48 already contained such a caveat, meaning only 66 contained a caveat inserted by us. Conducting the analysis as an intention-to-treat, there was no difference in the proportion of news articles that contained caveats between conditions (95% confidence interval of the odds ratio = .9 to 2.6).

For news uptake, there was no difference in the incidence of news reporting on press releases assigned to the control and statement of relationship group versus the two conditions that involved addition of caveats: 95% confidence interval of the odds ratio = .8 to 1.03.

4.4.3. As-treated – statements of relationship

This analysis used press releases in their final published form. In other words, the actual content of the final press releases was used for analysis, disregarding the condition to which each was assigned. This is similar to the type of analyses performed in chapters 2 and 3. News content was compared for press releases in which the headlines and main statements were aligned to the study design (correlational or equivalent statements for observational research, and causal or ‘can cause’ statements for experimental studies) versus press releases in which

the headline or main statements were not aligned to the study design. The odds of news containing aligned headlines was higher when the press release was aligned, versus when it was not: 95% confidence interval of the odds ratio = 1.6 to 5.0. The odds of news containing aligned statements of relationship in the top-line of the body of text was higher when the press release statements were aligned, versus when they were not: 95% confidence interval of the odds ratio = 1.8 to 4.8.

For news uptake, there was no difference in the incidence of news for press releases that contained aligned statements versus press releases that contained statements that were not aligned: 95% confidence interval of the odds ratio = .7 to 1.7.

4.4.4. As-treated – caveats

The odds of news containing a caveat statement was higher when the press release did, versus when it did not: 95% confidence interval of the odds ratio = 16 to 156. Almost no news articles contained caveats if the related press release did not.

For news uptake, there was a higher incidence of news for press releases that contained caveats, than press releases that did not: 95% confidence interval of the odds ratio = 1.3 to 2.7.

4.4.5. Condition non-adherence

As an example of the extent to which the content of press releases differed from that dictated by their assigned condition, for the statement of relationship analysis, the percentage of aligned claims in press release main statements was 91% (95% confidence interval = 82% to 96%) for those assigned to the intervention conditions, versus 82% (95% confidence interval = 77% to 86%) for those assigned to the control conditions.

For the caveat analysis, the percentage of press releases containing caveats in the intervention conditions was 40% (95% confidence interval = 30% to 51%) versus 17% (95% confidence interval = 7% to 36%) for the control conditions.

4.5. Discussion

News content and uptake was no different between conditions, but there was a high level of condition non-adherence whereby press releases frequently contained characteristics that were not assigned in their condition; for example control press releases containing both caveats and aligned headlines. Analysing the content of news based on the final content of press releases (similar to the analyses reported in chapters two and three) revealed a relationship between their levels of alignment and inclusion of caveats. Press releases with statements of relationship that were aligned to the study design also had no lower news uptake than misaligned press releases, in other words there was no evidence of a penalty for the alignment of statements. Counter intuitively, press releases that contained caveats to study design actually had higher news uptake. When press releases contained caveats the news was more likely to contain similar caveats.

The high level of condition non-adherence meant that a typical per-protocol analysis of the outcomes in news was not possible. This means that the remaining analyses are either confounded by the high level of condition mixing (intention-to-treat analysis), or to counter this issue were purely observational (as-treated analysis) based on the actual content of the press releases. This means that conclusions about the effect of press release content on news uptake and news content cannot be inferred. However, the intended changes in each experimental condition were positive changes; it would be a good outcome if press releases contained caveats and more aligned statements. The evidence points to spontaneous adoption of such content. This is in line with the findings of chapter three that press releases essentially became more accurate over time.

A difference between the analyses in chapters is the use of alignment rather than exaggeration. Exaggeration is one dimensional, in that it can only detect misalignment of statements to correlational designs. The findings of experimental designs cannot be exaggerated, and are thus not used in the analyses – limiting the size of the sample. Alignment on the other hand can include experimental study designs and this allows for the interpretation of the outcome to be with regards to overall accuracy of reporting regardless of design. Exaggeration is still a measure of accuracy, but a reduction in exaggeration is an increase in accuracy, whereas an increase in alignment is an increase in accuracy. The use of alignment as a homologue for accuracy is a better approach for future research, given that press offices are the primary stakeholders in this type of research, simply because the term is more suitable given its positive

connotations. InSciOut frequently attend press officer events to disseminate our research findings. At the STEMPRA Scotland event in May 2018, press officers showed a keen interest in our findings, and some expressed anxiety regarding the research. This is anecdotal, but the use of alignment can benefit relationships with press officers, and their interpretation of the findings as impetus to develop, and well as benefitting the research. The further experiments in this thesis will use the term exaggeration, since this is how they were designed.

Given that this analysis was observational, there is still a need for experimental investigation of the effect of press release content on news content and uptake. The trial was by no means a failure. Given that there was a high level of control in this research in comparison to previous retrospective observational designs, it could be argued that this trial represents stronger evidence of the relationship (Rubin, 2008) between press releases and news.

The findings overall point to a positive relationship between press releases and news, where more accurate press releases may lead to positive outcomes in news, although this needs supporting with experimental evidence. The lack of a penalty in news uptake, and the potential positive impact of caveat inclusion on news uptake for more accurate press releases is a promising result for the practice of press officers. Press officers know what to do to create more accurate, and arguably higher quality press releases, and in many cases they already are.

CHAPTER FIVE – EXPERIMENTS WITH JOURNALISTS

Note on contribution

Dr. Alice Rees assisted with the coding of participants' responses in the analysis of the effect of press release exaggeration on news content. Dr. Michael Hill, and Dr. Anne Harbin assisted with providing access to the journalism masters student cohorts for Cardiff University, and the University of West England, respectively.

5.1. Introduction

Previous chapters have focused on the relationship between exaggeration in news articles, and exaggeration in the content of source articles. The findings have been largely consistent with previous observational research (Schwartz, Woloshin, Andrews, & Stuckel, 2012, Sumner *et al.*, 2014; Sumner *et al.*, 2016, Schat *et al.*, 2018). The background research has been limited in that the observational nature of the studies limit our ability to draw definitive causal conclusions. In such studies, it is impossible to infer whether the effect would hold up for identical press releases that contained only manipulation to the variables of interest. In order to overcome this, and to test for this relationship using experimental methods, two studies were devised. The first was the randomised controlled trial outlined in chapter four, where real-world draft press releases were manipulated prior to release, and the subsequent content of news was monitored. The second study, described in this chapter, tested the content of articles written by journalism students in response to controlled experimental manipulations of the source articles, with counterbalancing to overcome the issues with the observational methods.

In addition to the discourse about the potential influence of press releases on the content of news, there is also interest in the potential influence of press release content on uptake of stories into the news. A key motivation for any press office would be to create press releases that spawn news content, and it may be that such pressures might allow exaggeration to creep into articles in an attempt to succeed. Exaggeration in press releases has both been shown to not be related to any change in news uptake (Sumner *et al.*, 2014) but also to be associated with an increased chance of news coverage (Schat, Bossema, Numans, Smeets, & Burger, 2018). The inclusion of caveats was also not related to any decrease in the uptake of the story into the news (Sumner *et al.*, 2016), contrary to the idea that inclusion of limitations could damage the reputation of a story. Of course, conclusions based on non-significant results must be cautious. An issue with the mentioned studies is that they are observational and retrospective, and cannot compare the effect of identical press release material with modifications to language and inclusion of caveats.

The present research was devised to experimentally examine whether manipulation to the language used in statements of relationship, and insertion of caveats of study design in real-world press releases can influence the perceived newsworthiness of the stories reported in the articles. The effect of exaggeration in press releases on the content of news stories written by participants was also tested. Masters level journalism students were provided with a selection of health-related press releases reporting on associative scientific research. These press releases were presented with either statements of relationship that were

representative of the journal article (associative, ambiguous, or conditional cause) or exaggerated statements (can cause, or cause), in accordance with the findings of Adams *et al.* (2017) either include or omit a caveat regarding the observational design of the study. Participants were asked to indicate which science stories were newsworthy, to give reasons for their selections, and to write stories. The hypotheses being tested were: 1) does the strength of language, and the inclusion of caveats influence participants' indication of the newsworthiness of health-related articles? 2) What are the factors that participants cite as reasons for their selections? 3) Does exaggeration in press releases influence the content of news articles?

5.2. Method

5.2.1. Participants

Twenty-nine students (16 female, age 21-29, mean = 23.9, SD = 2.21; 13 male, age 21-29, mean = 24.2, SD = 2.44) studying Masters degrees in journalism were recruited through their course coordinators at Cardiff University and the University of West England (UWE). The incentive for students was remuneration at a rate of £10 per hour. This cohort was selected because their courses had a significant practical element where journalistic practices were taught and simulated in a mock-newsroom environment. The mock newsrooms were essentially computer rooms equipped with facilities and equipment to facilitate news production (editing software, and sound recording booths for example). This meant that all participants had some form of experience acting out the different roles in a modern newsroom, and should be better suited to making judgments about newsworthiness of articles than participants from the

general population. Previous journalistic work experience was not a requirement of these courses, with students either being required to hold a higher education degree, equivalent qualification, or previous relevant work experience. Four testing sessions were held, three with separate cohorts of students from the University of West England (27 participants tested between 2016 and 2017), and one session with students from Cardiff University (2 participants in 2017). Ethical approval to test participants was obtained from the ethics committee in Cardiff University School of Psychology under project number (EC.16.03.08.4482), in agreement with the gatekeepers for each cohort of students.

5.2.2. Materials

In order to test the effect of exaggeration in press releases on participants' selections, and on the content of participants' news articles, a set of health-related press releases and their source journal articles were required. It was decided that rather than creating such materials, even though this would bring greater control to the experiments, the use of real-world press release/journal article pairs would present a more realistic task for participants. The press release and journal article pairs presented to participants were sourced from those collected by Sumner *et al.* (2014). Firstly, the corpus was searched for representative journal articles, that is, only cases where the journal article's main conclusions, abstract, and title all reported associative findings and where the study design was observational. From this selection eight exaggerated, and eight representative press releases were randomly drawn; that is, eight press releases where the strongest relationship stated in the title/main statements was

correlational or equivalent (those stating correlation, ambiguous, or conditional causal relationships); and eight where the title and/or main statement stated a relationship equivalent to ‘can cause’, or ‘cause’. In order to control for other differences (such as article topic) between these naturally exaggerated and representative press releases, modified copies were made where exaggeration was inserted (into copies of the representative press releases) or removed (from copies of the exaggerated press releases). For example, for press releases with exaggerated language, a representative version was created by inserting a modal verb into the headlines and main claims (e.g. inserting ‘may’ in “New antidepressants [may] increase risks for elderly”) or replacing the causal expression with an associative expression (e.g. replacing ‘increase’ with ‘associated with’ in “New antidepressants [~~increase~~] [associated with] risks in elderly”); for press releases with weak language, the modal verb was removed or associative expression replaced with a causal expression.

This produced a set of 16 press releases that could be presented in either exaggerated or representative formats. In addition to this, caveats regarding study design were then inserted into half of these press releases. The structure of caveats was formulaic, and mimicked those inserted into real-world press releases in Adams *et al.* (2019). To fit into the narrative of each press release, each caveat was bespoke but always contained 1) a mention of the study design, 2) a mention of the inability of this type of research to provide evidence for cause and effect, 3) a mention of the type of study design that can conclude cause and effect. For example, “As this was an observational study we cannot conclude that breastfeeding directly affects behaviour, other factors may have been involved

and would need to be investigated with an experimental study aimed specifically at uncovering cause and effect”, or “This observational study contributes to the evidence showing that exposure to family violence is related to brain function, but we cannot rule out other factors with this type of research, we cannot make conclusions about cause and effect - for that, we would need to conduct an experimental trial.”.

This process meant that the 16 original press releases could be presented to participants in any combination, with or without exaggeration, and with or without caveats.

5.2.3. Design

5.2.3.1. News selection

To assess the effect of press release content on news selection, each participant was provided with one version of each article. A 2x2 within-subjects design was employed. The two factors (exaggeration, and caveat) were manipulated within subject and within item, giving four conditions (table 5.1). The assignment of items to conditions was counter-balanced so that no single participant saw the same press release twice, but across participants, all press releases occurred in all conditions. An implementation error with counterbalancing was detected after the experiment was completed. One group of participants saw five items in the caveat condition and three in the non-caveat condition, and another saw three in the caveat condition and five in the non-caveat condition.

Table 5.1.
Experimental conditions for the news selection and news content experiments.

News selection		News content	
Representative language	Exaggerated language	Representative language	Exaggerated language
Caveat	Caveat		
Representative language	Exaggerated language		
No caveat	No caveat		

Each participant was presented with 20 press releases/journal article pairs in total: the 16 press releases detailed above, and four filler items in which the press release and journal article both contained exaggerated statements. These items were included for comparison, to assist with interpretation of the findings in the event that participants select more exaggerated articles than representative articles. Given that the exaggeration condition only differs in that press releases contain causal statements, it would be difficult to interpret whether it is exaggeration that makes the articles more attractive, or if it is merely that causal statements are more attractive. The filler items contain causal statements, but they do not exaggerate the statements presented in the journal article. Thus, if the selection of exaggerated press releases is elevated, but the selection of filler items is not, then it could be concluded that it is exaggeration that makes a news article more attractive.

Each participant gave a binary decision for each press release: newsworthy or not newsworthy. The outcome measure was the proportion of articles selected as newsworthy in each of the four conditions.

5.2.3.2 News content

To assess the effect of exaggeration in press releases on the content of participants' news articles, four of the original 16 press release/journal article sets were used. Exaggeration in the articles was manipulated within subject and within item, such that each press release was presented in both exaggerated and non exaggerated formats between participants, and each participant was presented with one exaggerated and one representative press release. The articles that participants saw in this phase were presented in the same condition as the news selection experiment, so that participants did not see the same article with two different statements or relationship between experiments.

Two independent coders assessed participants' articles using a coding sheet adapted from that used in chapter two (an example of the updated coding sheet is provided in appendix 5.1). This sheet recorded the statements of relationship written by participants in the main text, and if present, the title. The highest code (i.e. strongest relationship statement) of the two was then taken as the statement to be used for analysis. During coding, information about the conditions of the experiment was obscured so that it could not influence the process. Both coders' interpretations were then compared, and differences were resolved through consensus agreement.

The definition for exaggeration was based on the findings of Adams *et al.* (2017), where correlational, ambiguous, and conditional causal relationships (for example, may/might/could cause) were interpreted to be equivalent, and relationships equivalent to 'can cause' or 'cause' were interpreted to be of higher

strength. Since the materials used in the present study are based on correlation or equivalent findings, participant stories that contain relationships equivalent to ‘can cause’ or ‘cause’ will be defined as exaggerated. Thus, the outcome measure is the proportion of participants’ news stories that contain exaggerated statements, across the two conditions (presented in table 5.1).

5.2.4. Procedure

The experiments were conducted in a mock newsroom. This was an attempt to emulate the centralised newsroom environment in which modern journalists work (Williams & Clifford, 2009) and bring greater validity to the experiment. Each participant was provided with a storage drive containing a set of press release and journal article pairs in folders according to the experiment. In a continued effort to maintain validity, participants were told that they could use any resources they would normally use to make selections (for example, using internet dictionaries and encyclopedias to define terminology), but in the interest of time, they should not attempt to contact sources or experts for further information (contrary to their normal strategy). They were also not allowed to communicate with other participants so that they would not be exposed to alternate versions of their materials.

For the selection experiment, participants were given 40 minutes in which they should indicate whether they believed the scientific findings reported in each press release-article pair were newsworthy by indicating “yes” or “no” to the instruction “please indicate whether you think this research should be put forward for a news article”. Participants were not told to make a particular

number of selections but most selected around half as newsworthy, with an average of 55% of press releases selected per participant. In addition to indicating the newsworthiness of each story, participants were asked to provide a reason for their decision. Participants were given the instruction “Please give a brief reason/s for your decision in the box below” with a text box for a free-text response.

In the news content experiment, participants were given 40 minutes to write two news articles, one for each of the press release/journal article pairs presented to them. They were instructed to write news-in-brief articles, which are typically around 250 words in length.

To further assist with interpretation of the results, participants were also asked to respond to two questions, should they have time after completing the experiments. The two questions were: 1) ‘To what extent did you use/find useful the original journal articles?’ and 2) ‘Did you use/find useful any sources other than those provided? If so, which? (e.g. other news websites, journal webpages etc.)’. Question one was presented to understand whether the content of the journal article is taken into consideration when making selections, and should assist with interpretation for the same reason as the inclusion of the causal, but not exaggerated, filler phrases. Question two was included to help understand whether participants felt enough information is provided in the press release, or whether they looked for information elsewhere. Given that the press releases and journal articles were taken from a real world sample, this question was also

intended to help identify cases, if any, where participants became exposed to the online version of the modified article they had been presented with.

5.3. Results

5.3.1. News selection

A binomial logistic regression compared the effects of press release language exaggeration, inclusion of caveats, and the interaction of language and caveats, on participants' selections. Prior to this step, the press release items were entered into the model to check for effects of other between-item characteristics. That is, the press releases grouped by the original article, not differentiating between modified or natural press releases. Participants could not be shown the same article twice, since this would reveal the experimental manipulations, so this step is necessary to test for variation between press releases. For this step, the press releases were each turned into dummy variables and the effect of each was compared to press release 14-11-18 (antidepressant use in elderly patients). This was an arbitrary selection, merely to test for variation in participants' selections between press releases. The model was significant in the first step $\chi^2(15) = 52.317, p < .05$ (Table 5.2), explaining 14.4% (Nagelkerke R^2) of the variance in press release selection, and correctly predicted 64.8% of selections. Five press releases (02-11-027, 03-11-010, 19-11-011, 06-11-015, and 07-11-040) had significantly lower odds of being selected by participants: inverse odds ratios = 10.10 (95% confidence interval = 2.99 to 33.33), 4.72 (1.48 to 14.93), 10.10 (2.99 to 33.333), 12.05 (3.50 to 41.67), and 4.72 (1.48 to 14.93) respectively.

Table 5.2.

*Inverse odds ratios and 95% confidence intervals for press release items as predictors of article selection. Press release 14-11-18 was used as an indicator variable for comparison. * $p < .05$*

Press release	Lower 95% CI	Odds ratio	Upper 95% CI
14-11-018 – antidepressants (indicator)			
02-11-028 – pregnancy blood composition	0.89	2.87	9.26
06-11-012 – bowel cancer	0.98	3.12	9.90
15-11-011 – breastfeeding	0.65	2.13	6.94
03-11-026 – chocolate and heart health	0.73	2.34	7.58
02-11-027 – genetics and arthritis	2.99	10.10 *	33.33
03-11-010 – parent and child happiness	1.48	4.72 *	14.93
19-11-011 – heart disease	2.99	10.10 *	33.33
07-11-043 – iron and blood clots	0.52	1.72	5.68
02-11-018 – jogging	0.89	2.87	9.26
19-11-014 – maltreated children	0.28	1.00	3.56
14-11-006 – morning-after pill	0.62	2.02	6.58
06-11-015 – obesity	3.50	12.05 *	41.67
07-11-040 – premature birth	1.48	4.72 *	14.93
05-11-033 – schizophrenia	0.89	2.87	9.26
08-11-022 – vitamins and asthma	0.85	2.70	8.70

The strength of language, inclusion of caveats, and the interaction of language and caveats were then entered into the model. With press release items controlled for, this step made no statistically significant contribution to the model $\chi^2(3) = 1.125, p = .77$ (figure 5.1) with neither exaggerated language (odds ratio .82, 95% confidence interval .47 to 1.42), inclusion of caveat (.87, .50 to 1.51), nor the interaction of the two (1.50, .69 to 3.26), altering the odds of press release selection.

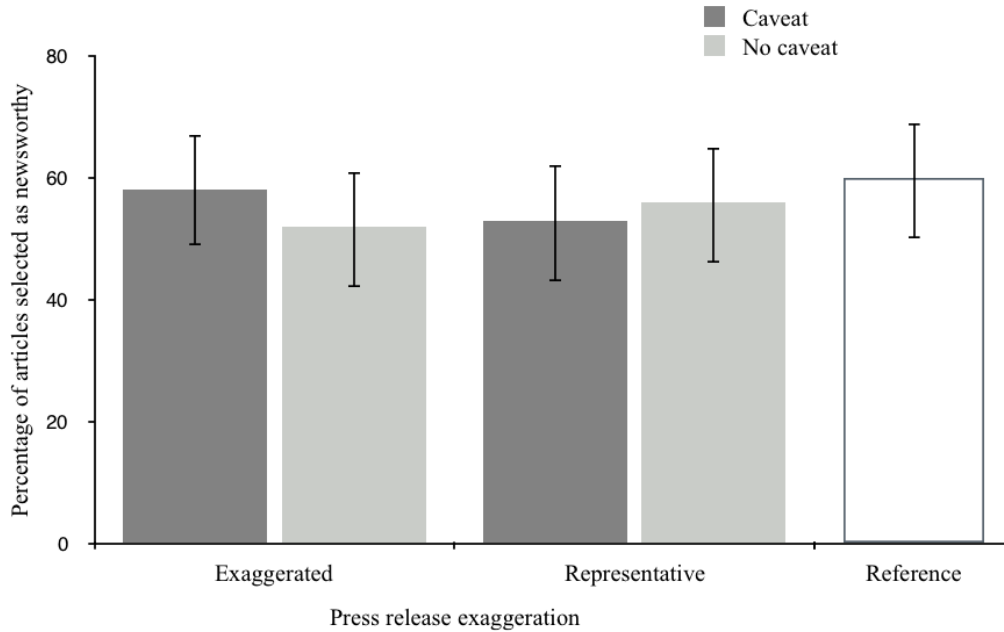


Figure 5.1. Percentage of press releases selected as newsworthy as a function of press release exaggeration and inclusion of caveats. The white bar indicates selections for articles where both the press releases and journal articles exaggerated. Error bars are bootstrapped 95% confidence intervals.

5.3.2. Analysis of comments – reasons for news selection decisions

Participants provided 394 free text comments outlining their reason/s why they selected or rejected each press release as newsworthy. In order to interpret the comments, the first three phases of thematic analysis as outlined by Braun and Clarke (2006) were followed to identify semantic themes. Since this is an exploratory analysis, an inductive approach was used whereby comments were interpreted in a bottom-up fashion, without reference to any theoretical background during coding. As with any thematic analysis, it has to be noted that the coder's background and theoretical understanding are likely to be a source of variation in the interpretation of themes, but the inductive approach should be largely robust. From the 394 comments, 628 individual reasons were coded. Of

these, 267 reasons were given for not selecting a press release, and 361 reasons were given for selecting a press release. These reasons were reduced to 28 distinct categories. A comparison of the frequencies of responses is given in Table 5.3. Most of the categories were given as reasons both for and against selection. For example, some participants did not select articles because further research was needed, but participants also selected articles because they justified further research. These two instances were coded as one category regarding future research. Articles that were not selected because they appeared misleading, and articles that were selected because their subject matter is generally misreported were combined.

Table 5.3

Frequencies of reasons given for press release selections and non-selections. Reasons are ranked in descending order of frequency given for selections and non-selections combined. Each reason can be given as a basis for selection or non-selection.

Rank - Reason	Not selected	Selected
1 - appealing/interesting	32	55
2 - size of audience	17	55
3 - novelty	26	22
4 - complexity	31	10
5 - helpfulness to reader	4	27
6 - impact of research/implications	2	28
7 - study quality	15	15
8 - specific target (e.g. specific audience)	20	5
9 – topic (popular, current, over-reported)	4	20
10 - further research needed/justifies further research	15	9
11 - influence behaviour of readers	9	14
12 - relationship strength	17	6
13 - importance	6	16
14 - common knowledge	18	2
15 - attention grabbing (“groundbreaking”)	9	9
16 - negative (including controversial)	14	3
17 - shareable (social media)	0	16
18 - entertaining	1	15
19 - accessibility (ease of understanding)	2	12
20 - misleading (a topic generally reported misleadingly)	11	2
21 - source quality (e.g. "Cambridge University")	0	9
22 – press release quality	4	1
23 - caveat (inserted caveat only)	5	0
24 - debatable (sparks discussion)	0	4
25 - human interest	3	1
26 - positive	0	3
27 - political	1	2
28 – balanced reporting	1	0

Given the disparity in the number of reasons given between selected and rejected press releases, care must be taken when interpreting the frequency of occurrence of themes. Relationship strength, and the inclusion of caveats (the experimental manipulations) did emerge as themes, though infrequently. There were 23 instances where participants specifically mentioned the strength of relationship (category 12, ‘relationship strength’) as a reason for selection (6 cases), or for rejection (17 cases). There were five instances where participants cited the

inserted caveat (category 23, 'caveat - inserted caveat only'), for example one participant wrote "They also say 'As this was an observational study we cannot conclude that breastfeeding directly affects behaviour'. So I feel that there are still more research to be done before this becomes newsworthy", directly quoting the inserted caveat. The inserted caveat was only cited as a reason to reject a press release. Furthermore, the 'further research needed' (category 10) may also be relevant to the caveat manipulation since this reason to reject a press release was mainly given in response to press releases with inserted caveats. Examples of responses in this category, such as "If more research needs to be done it is not conclusive yet and not groundbreaking enough to make it newsworthy", and "The story is that researchers are one step closer to fully understanding this illness, not that it has been cured or a new treatment has been discovered." imply that the research cannot infer causation, and more research is needed.

It is important to note, that although reference to the manipulations appeared, they did so infrequently. Other factors relating to the press releases were far more prominent, such as whether the story was 'appealing/interesting' (category rank 1), the size of the potential audience (category 2), the novelty (category 3) of the research, and its complexity (category 4).

5.3.3. News content

A binomial logistic regression was used to analyse the occurrence of exaggeration in participants' news stories, in comparison to the exaggeration in press releases. In the first step, the press release items were entered into the model, and in the second step press release exaggeration was entered. The model was not significant in the first step $\chi^2(3) = 1.988, p = .58$, meaning that press

release item did not significantly predict variation in participants' article content. The second step, where press release exaggeration was entered into the model, was also not significant $\chi^2(3) = 1.619, p = .20$, meaning that press release exaggeration did not significantly predict variation in exaggeration in participants' articles. Figure 5.2 demonstrates the levels of exaggeration in participants' articles.

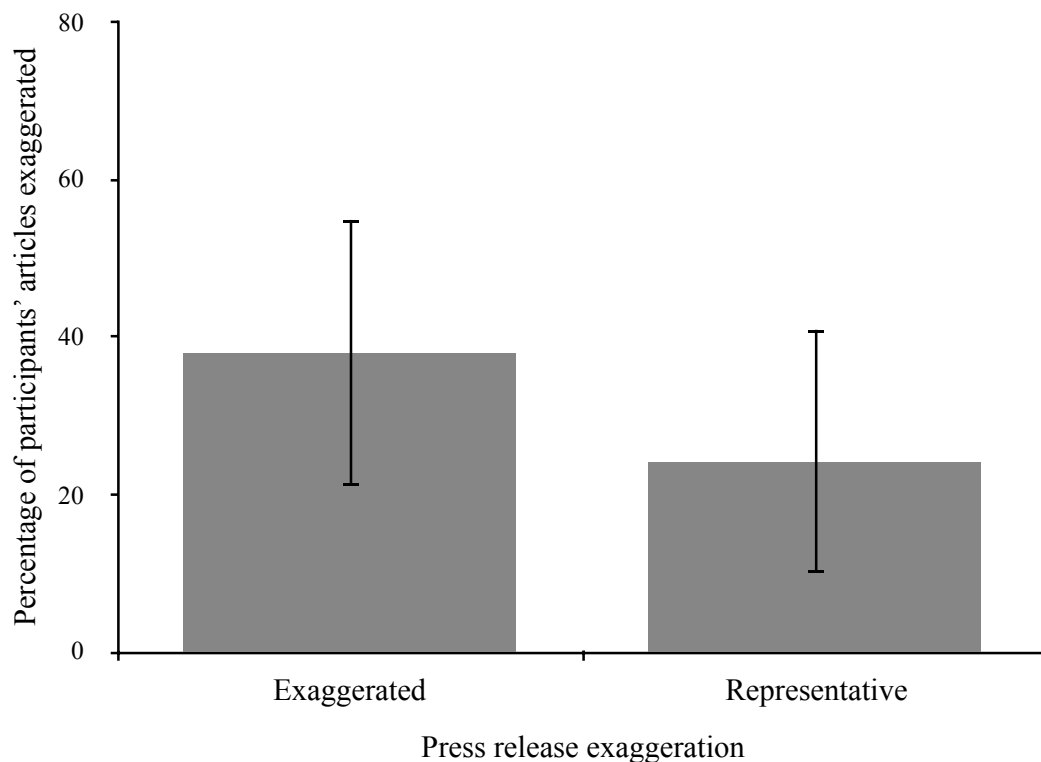


Figure 5.2. Percentage of participants' articles containing exaggeration as a function of press release exaggeration.

5.3.4. Other sources used

There were 16 free text responses to the question “Did you use/find useful any sources other than those provided? If so, which? (e.g. other news websites, journal webpages etc.)” (provided in appendix 5.2). Most frequently, participants referred to the time constraint preventing them from using other sources, or from spending enough time with other sources. Two participants who reported the

time constraint also mentioned how the press release was sufficient to write their article: “I didn’t look anywhere else, partly because of the time constraints. Also usually a press release and the report is enough anyway”, and, in reference to other sources: “Under the time constraints, I did not use them. All information was in the press release”. Other sources mentioned were generally used to define terminology, or for a quick clarification. For example, one participant described checking for an acronym “I checked the acronym for Department of Public Health and Primary Care, that was it”.

There were 17 responses to the question “To what extent did you use/find useful the original journal articles?” (provided in appendix 5.2). In 12 of the 17 responses, the time constraint was again mentioned as a reason for not making use of the journal article. For example, one response succinctly stated: “If I had more time I would have used them more”. One participant expanded on this issue by interpreting the lack of time as “worrying, because it means that whoever writes the press release has far too much control”. One participant was conscious of the lack of time preventing them from referring to the journal articles in the news selection experiment, so they attempted to use the journal articles in the content creation experiment because “there is danger in my lack of scientific background leading me to misinterpret the information, I nonetheless feel it behoves me to look – as the press releases alone may be misleading, in either overstating their findings or oversimplifying the data”. Participants also frequently referred to the complex, specialist, or confusing nature of the journal articles as a reason to avoid them. With reference to the news selection experiment, one participant implied that they should not need to refer to the

journal article because “if it can’t be easily summed up in a press release then it is unlikely to be a story worth covering in a short amount of time”.

5.4. Discussion

The present research was designed to overcome the limitations of retrospective observational methods, to create an adequate control between items, and to manipulate variables directly. The investigation set out to determine if the presence of causal language and caveats could predict participants’ indication of the newsworthiness of health-related articles, and to investigate the factors participants cite as their reasons for making their selections. The second phase of the experiment aimed to measure whether manipulating exaggeration in source material would influence the content of participants’ news articles.

5.4.1 News selection

Though the previous literature demonstrates examples of associations between the characteristics of press releases and news articles (Sumner *et al.*, 2014; Schwartz, Woloshin, Andrews, & Stukel, 2012; Schat, Bossema, Numans, Smeets, & Burger, 2018), there appears to be no such relationship, or no detected relationship, between press release characteristics and uptake into news. A true absence of an effect of the manipulations in this study would suggest that although caveats and language are transferred between the stages of the science media process, they might not be particularly important in the decision making process of journalists and editors who are looking for an article to sell to their readers. Further to this, cases where press releases and journal articles both contained exaggeration (the reference - or ‘filler’ - items) appeared to be as equally

likely to be selected for news as the experimentally manipulated press releases (based on bootstrapped 95% confidence intervals), meaning that exaggerated articles may not be discernable from any articles containing causal language. However, this interpretation is used with caution in the absence of an effect. The finding that the 16 different press releases items were a relatively good predictor of variation in participants' selections was not predicted. They were entered into the model to control for variation before the manipulated variables were entered into the regression model. Since there was no prior reason to assume that any press release would influence uptake more than any other, the selection of an indicator was arbitrary. Coincidentally, the press release selected to be the indicator variable (antidepressant use in elderly patients) was also most frequently selected by participants to be newsworthy. It cannot be certain which characteristics about the press releases made some more predictive of selection than others. The press releases were selected to be matched on many levels, such as the design of the research being observational, the language being classified in accordance with the findings of Adams *et al.* (2017), and the structure following an inverted triangle approach with the statement of relationship in the title and first lines of text. However, the articles did vary on some other measures not expected to have an effect (and thus not quantified in the described experiment), such as word count, inclusion of postscript, and institution of origin. Press releases also varied on topic (e.g. health benefits of chocolate, or effect of exercise on bone health), and sample population (e.g. the elderly, pregnant women, males residing Scotland). It intuitively makes sense that some topics will be more appealing than others, or some samples might be more widespread or

important in the population. Indeed, the most frequent reason given for selecting a press release was “appeal/interestingness”.

Through it is interesting which topics are more appealing to an audience, it does not lead to an important implication for scientists, press officers, and journalists. The topic of published research cannot be, and should not be attempted to be, manipulated in press releases and news with an objective to alter public appeal, and to gain more coverage. If the inclusion of caveats and alteration to language does not alter perceived newsworthiness, this does not mean that they are not important aspects for science writers to take into account. We should still follow the best practices of guidelines such as those published by the Science Media Centre (2012), and the Academy of Medical Sciences (2017), in providing accurate and honest information so that this may have a higher chance of being transmitted to the general reader to better inform health-related behaviour. Though as demonstrated in the news content experiment discussed below, this transfer of information to the news does not appear to hold up under experimental scrutiny.

Given the large variation recorded across press release type, future research could benefit from using an alternative approach to designing the materials. In this study, real world press releases were modified and given to participants, such that each press release was presented in every condition, but this also meant that other variations between the articles existed. Future research could remedy this issue by preparing a set of formulaic press releases, such that as many characteristics are controlled between press releases. For example, rather than

using variables that the participants will be familiar with (such as smoking and breast cancer), the variables could be fabricated drug and disease names (such as Metazophine and H41-BT) which are presented to the participants under the guise of being new treatments for recently discovered diseases. This of course would bring limitations, given that participants will not be able to investigate the story using other sources, and the formulaic nature of each press release seen may reduce the validity of the experiment.

5.4.2 News content

In previous literature, news content has generally been found to be related to the content of the press releases on which they are based. This experimental approach did not reach the same conclusion. Though the trend was for more exaggerated articles to be written by participants when they were exposed to exaggerated press releases, this outcome was not significant. Given the extent of this difference reported in the observational literature (a difference of up to 86% between the exaggerated and non-exaggerated conditions, according to Schat *et al.*, 2018) a similar effect size would have been expected here, should this experiment be an valid representation of journalistic practice. A major difference between the observational literature and this experiment is that there is a higher level of control and a reduction in the variation that would be due to other variables present in the real world. For example, participants were heavily restricted from being able to use other sources, as their comments revealed. But a reduction in variation would be expected to magnify the effect of interest rather than diminish it. As postulated previously in this thesis, communication between scientists, press officers, and journalists, and the independent investigation by

journalists are extraneous variables that were not accounted for in the observational literature. In this study, there was no such communication as it was disallowed, and although participants were allowed to use the Internet for further research if needed, participants either did not attempt to do so, did not find it useful to do so, or simply did not feel they had the time.

Evidence from the surveys and interviews conducted by Williams and Clifford (2009) demonstrates the similarities rather than differences between the experiences of journalists, and the experiences of the present study's participants. Around 68% of experienced career-journalists surveyed either agreed, or strongly agreed that they did not have enough time to adequately research their stories. One journalist stated that the time they spend on covering articles from the journal *Nature* for a feature article (a longer article than the news-in-brief articles that the present study's participants were expected to write) had been reduced to around one hour in recent years. It is also unlikely that the time constraint was to blame because the feature of interest in participants' articles was the statement of relationship, which is typically found, and was found, in the first few sentences of news articles due to the inverted pyramid writing style.

The students do differ from career-journalists in two important ways. Firstly, given the early phase of their careers, and given that they are enrolled on the same course, they are likely to not hold, or not exhibit any particular affiliation to a level of the media market. A journalist working for mid-market newspapers like the *Daily Mail*, would be expected to write articles differently to those writing for the *Guardian*, or for red-tops such as *The Sun*. Participants in the

present experiment are likely to write articles in a style that is more closely aligned to the academic nature of their course, though it is unclear what style of paper this would most closely resemble. Secondly, in the modern centralised newsrooms, editors each operate according to their specific agenda, or the agenda of the newsroom, and heavily direct journalists, to the extent they often dictate the top line or instruct the direction of articles (Williams & Clifford, 2009). This layer of journalism was not present in the experiment. Participants were all given the same instruction to write a news-in-brief article on the topic of the articles provided to them with no reference to accuracy, audience, or justification for writing. It may be that the articles created by participants in this experiment were more closely aligned to an academic writing style given that they were written at university, which may be more accurate, and therefore cautious with regard to exaggeration.

A major limitation of this study was the difficulty in recruiting journalism students during the course of their studies. During the formulation of this study, it took a great deal of effort to gain access to both the participants, and the newsroom environments, given the tight schedules of the students, and the exclusivity of the facilities. It took two years to test the 29 participants from whom the data reported in this chapter were gained, meaning that the resources required to obtain a larger sample were beyond the constraints of this project. Given the potential differences between journalism students and professional journalists with affiliations to news corporations, it may be more efficient for future studies to collect samples from undergraduate students, or members of the general population who have experience writing articles.

CHAPTER 6 – INSTRUCTION EXPERIMENT

6.1. Introduction

Previous chapters have investigated press release characteristics and their influence on uptake and content. Experiments in chapter five into the selection of news and the content written by participants showed that there was a great deal of variation in their decision making related to other factors in the articles. It is known that the predispositions of journalists can affect their interpretation of a topic (Starck & Soloski, 1977). In a survey of journalists in the UK, 33% said their own opinion of the issue influenced their decisions on how to interpret news articles (Patterson & Donsbach, 1996).

Rarely though do journalists act completely independently. The vast majority are affiliated with news corporations. Even the expression of personal beliefs and opinions through news is facilitated by via the opinion sections of newsrooms. Another important factor to consider is the influence of the newsroom on the decisions made by journalists. Institutional objectives are one of the four main factors traditionally held in the field of communication in journalism as most influential over journalists' decision-making (Donsbach, 2004). This refers to the pressure from the values of the institution, or the editors' personal requirements. The alignment of news content in this way to the requirements of the organisation, or the conditions set by the editor is often referred to as 'slant'. Given the political nature of news reporting, slant is likely to vary between news organisations. In response to the question "How often is the news you prepare changed by another person in the newsroom?" 28% of journalist respondents in

the UK indicated that their content was often or occasionally changed to increase audience interest, and 13% indicated that changes happened to improve factual accuracy (Patterson & Donsbach, 1996).

Similar findings have been seen in experimental psychology. An investigation into language style matching with 1744 university students in the US compared the writing style of participants in response to four essay questions written in vastly different styles: scientific, informal, convoluted, or arrogant. The style of participants' responses was found to frequently mimic the style of the question they were given (Ireland & Pennebaker, 2010).

So it may be that variation in the content of science news articles may be explained better by the instruction provided to participants, rather than the actual content of the press release. In the experiments in chapter five, all participants received the exact same instruction, since all participants were tested at the same time in the same newsroom. The following study aims to examine the content of news articles based on health-related press releases in response to specific instructions to write the article in either a way that is more concise, a way that is more appealing to the general reader, or accurately. To avoid a low sample size, this experiment was conducted online with undergraduate psychology students. It is thought that the instruction given to participants will affect the content of their news stories. Instruction groups will be compared for the proportion of articles that contain exaggerated statements of relationship. The instruction to write concisely was included to see if exaggerated statements of relationship would creep into articles simply because causal statements are typically shorter than

correlational statements. The highest proportion of exaggerated statements are expected in the group who are asked to make the article more appealing, with the instruction to write accurately yielding the lowest exaggeration rates.

6.2. Method

6.2.1. Participants

One hundred and fifty two undergraduate psychology students (131 female, mean age = 19.7 years) were recruited through an online tool incentivised by course credit for participation in studies within Cardiff University School of Psychology. The participants were granted credit upon their participation in the study.

6.2.2. Materials

Eight real health-related press releases were used as the participant stimuli. Press releases covered a range of topics: geography/autism, premature birth/school performance, bowel cancer/wealth, diet/blood pressure, sedentary lifestyle/diabetes, smoking/birth defects, iron intake/birth-weight, and breastfeeding/child behaviour. These were the eight representative press releases randomly selected from the corpus of Sumner *et al.* (2014) for the experiments in chapter 5. All press releases were based on journal articles with observational research methods and made observational conclusions; in other words, they were free of exaggeration to main statements of relationship.

There were three possible instructions: 1) Your task is to reproduce the content of this press release in your own words; 2) Your task is to rewrite this press

release in a way that is more appealing to the general reader; or 3) Your task is to rewrite this press release in a way that is more concise.

The study was hosted on Qualtrics, a web tool commonly used for the creation of questionnaires and simple online studies. Ethical approval to test participants was obtained from the ethics committee in Cardiff University School of Psychology under project number (EC.17.08.08.4938).

6.2.3. Procedure

The study was advertised through the School of Psychology's course credit system where participants have a choice between active studies. Upon agreeing to take part in this study, participants were directed to the Qualtrics form. Firstly they were presented with the consent form. Clicking to advance past the consent form started the study. A withdrawal after this point, or a failure to finish the study was deemed as a withdrawal, and participants were fully awarded course credit. Participants who did not advance past the consent form were deemed to have disagreed to consent to the study, and were not allocated course credit.

After the consent form, participants were asked to download an image of one of the randomly allocated press releases. An image was provided, rather than a document file in order to discourage the participants from copying directly from the press release. After downloading the press release, participants were presented with a page displaying their randomly allocated instruction and a free-text box where participants could write their response. They were instructed to take no more than 10 minutes to write their article.

Upon completing the article, participants were presented with a debrief form, and were allocated course credit.

6.2.4. Design

A 3-factor-between subjects design was employed. The three factors were the concise, accurate, and appealing writing styles. Participants contributed one article to one condition. Assignment of conditions was randomised, leaving 51 participants in the ‘accurate’ condition, 50 in the ‘appealing’ condition, and 51 in the ‘concise’ condition.

6.2.5. Analysis

Cases where participants’ articles did not report the same variables as the press releases allocated to them were excluded. Two participants were excluded from the ‘appealing’ condition based on this criterion, leaving 150 participants responses for analysis. The resulting articles were anonymised before two independent researchers interpreted their relationship statements. The two coders then met to discuss any inconsistencies between their interpretations, and to come to a final consensus agreement. A third coder was consulted if any inconsistencies could not be resolved. Prior to this consensus agreement, there was a 90.2% rate of agreement between coders. The relationship statements were coded according to the scheme in table 6.1, where the relationship between variables is placed into one of the five categories. Articles rated as equivalent to ‘can cause’, or ‘cause’ would be deemed as an exaggeration of the press release. This provided binary ratings of inflation for each participant’s article. A binary

logistic regression was performed comparing incidence of exaggerated claims across the three instruction categories.

Table 6.1

Statements of relationship classification designed by Sumner et al. (2014) and developed by Adams et al. (2017). Categories are presented in order of increasing strength. Press releases used as stimuli were graded as category two.

Statement of relationship categories				
0 - No cause stated	1 - Statement of no relationship	2 - Correlation, ambiguous or conditional cause.	3 - Can cause	4 - Direct cause stated

6.3. Results

A manipulation check was performed to see if participants were responding differently between conditions. The average word count for articles in each condition was compared to see if participants assigned the instruction to write concisely actually wrote shorter articles. The Levene statistic of homogeneity of variance demonstrated that the variances in word count between instruction conditions was unequal, $F(2, 147) = 3.87, p = .023$, meaning that an ANOVA could not be performed. The median word counts between conditions was 140 (range = 48 to 439), 135 (range = 38 to 464), and 125 (range = 43 to 259) for accurate, appealing, and concise instructions, respectively. A Kruskal-Wallis test demonstrated that there was no significant difference in average word count between groups $\chi^2(2) = 4.48, p > .05$.

Across all instruction conditions, 24.7% of participants' articles contained exaggeration of the main statement of the press release presented to them. There were 51 responses in each of the 'accurate' and 'concise' instruction conditions, and 48 responses in the 'appealing' condition. A binary logistic regression was used, and data met the assumptions of a dichotomous outcome with mutually exclusive and exhaustive categories, independence of observations, and no continuous predictors were used. In the first step of the binary logistic regression, the press release items were entered into the model to account for variation in exaggeration across press release. Press release categories were coded as dummy variables and compared to the press release reporting on sedentary lifestyle/diabetes (an arbitrary selection). The model was significant in the first step $\chi^2(7) = 39.021, p < .05$, explaining 34.0% (Nagelkerke R^2) of the variance in participants' exaggeration, correctly predicting 77.3% of exaggerated cases. One press release, reporting on smoking/birth defects, had significantly higher odds of leading to exaggerated articles (odds ratio 5.19, 95% confidence interval = 1.23 to 21.56), with 57.9% of articles containing exaggerated statements of relationship. Two press releases (geography/autism, and premature birth/school performance) were never exaggerated by participants. In the second step, instruction category was entered into the model. Categories were dummy coded and compared to the instruction to 'reproduce the press release accurately'. This step was not significant $\chi^2(2) = 3.345, p = .188$. Figure 6.1 demonstrates the rates of exaggeration across the three instruction groups.

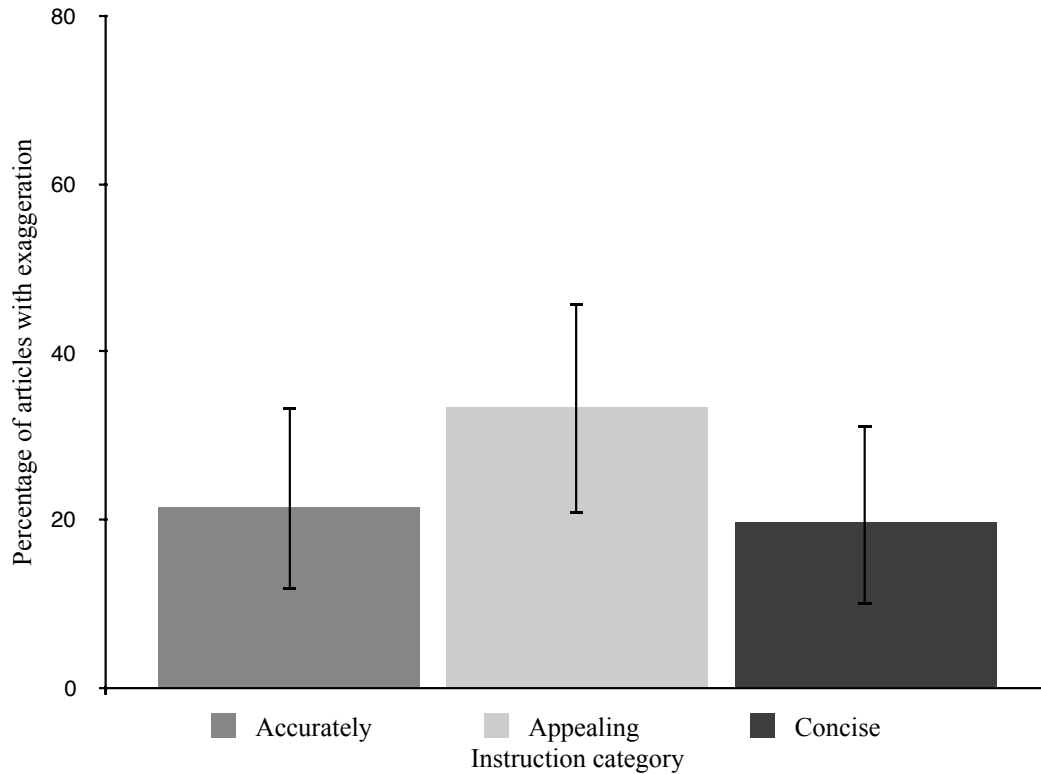


Figure 6.1. Percentage of articles containing exaggeration across the three instruction conditions. The 'accurately' condition was used as the indicator variable.

6.4. Discussion

This study aimed to discover if giving instruction on the approach to interpreting a health-news story would influence the presence of exaggeration in participants' articles. The three instructions, to rewrite the article accurately, in a more appealing, or a more concise way did not explain variation in the levels of exaggeration included in participants' articles. Entering the press release item into the regression model to assess variation in exaggeration between press release items significantly explained the variation. Exaggeration in participants' articles across press release item varied from 0% to 57.9%, meaning that characteristics other than the writing objective are likely more important in

explaining exaggeration.

The lack of an effect of instruction cannot be interpreted adequately with this analysis, since there was no statistical difference in exaggeration between groups. But if this is indicative of a genuine lack of a difference, it is curious to consider that the instruction to reproduce the story accurately might yield stories that are no more accurate than stories created under the instruction to sensationalise. Perhaps the homogeneity of the psychology undergraduate sample nullifies the predicted difference between groups given that they have a more similar background to each other than journalists from different newsrooms have with each other. Although the overall level of exaggeration in participants' articles is approximately equal to the levels of exaggeration found in real-world news articles reporting on representative press releases from 2014 and 2015 (see chapter 3). This suggests that this participant sample may demonstrate behaviours that are largely representative of those exhibited by journalists, despite this participant sample originally being seen as a compromise of journalistic experience in favour of a larger sample size.

It was thought that a higher level of exaggeration in the group that were explicitly told to make the article more appealing would indicate that participants see exaggeration to statements of relationship as an appealing modification. It could be however that the instruction did have an effect on another aspect of the articles, but in the interests of saving time in the coding process for this study, only exaggeration in statements of relationship was coded as an outcome variable.

The manipulation check comparing word counts across articles was intended to discover whether the instruction conditions were changing participants' behaviour. The lack of a difference in average word count between groups suggests that either participants were not responding differently between conditions, or that potentially that participants were not given enough time to type the longer articles that would be expected to occur in the accurate and appealing conditions. If this finding was indicative of a failure of the manipulation to alter participants' behaviour, this would support the lack of a significant difference in exaggeration between groups. In this case, it may be that the instructions to write in different styles was misunderstood, or simply that participants may have little to no prior experience of attempting to write in different styles. On the other hand, participants were only given 10 minutes to write their articles. At a rate of 30 words per minute, we would expect to see an average word count of 300 words. None of the participants in the 'concise' condition wrote articles exceeding 259 words, whereas seven participants in the other conditions exceeded that word count, with a maximum word count of 439 in the 'appealing' condition. Assuming that between groups there was an equal proportion of participants who had the ability to type more than 259 words in the time scale, the question raised is whether those participants had finished their 'concise' articles, whilst there were unfinished articles in the other conditions that would have been longer given more time. There is unfortunately no way to check the articles to see whether they were complete because news articles following the inverted-pyramid style do not have an obvious conclusion. In future, it may be wise to give the participants more time, ask them whether they completed their article or needed more time, or simply by giving them the

objective of finishing their article, rather than a time limit. This would increase the chance that a true effect of the manipulation would be found should it exist. A time limit was of course initially chosen to better replicate the pressures of the newsroom, so removing this might impact the validity of the study.

The undergraduate students employed as participants in this study will also have a good level of understanding when it comes to interpreting and writing scientific statements, due to the content of their course, and their frequent practical report writing assignments. So it could be that such participants are likely to employ a certain scientific style regardless of the instruction given, which may always yield approximately the same output. However, this study only examined the relationship statements themselves, it could be that any difference between the instruction conditions could manifest in other aspects of their writing which may only be picked up with qualitative analyses such as that employed by Ireland and Pennebaker (2010).

The unexpected finding that there was a large variation in exaggerated statements between the press releases items suggests that there must be some other aspect of the press releases that influence participants. For future studies, a solution to the issue of differences between press releases could be to create formulaic press releases where content (word count, topic, structure) is controlled. For example, the articles could report on relationships between fabricated and anonymous drug and virus combinations as the variables of interest, such that the articles will be novel to all participants. This could also avoid any variation in responses due to the predispositions of participants such as those reported by Starck and Soloski

(1977); and Patterson and Donsbach (1996).

In conclusion, preconditions for writing style may not influence the inclusion of exaggeration in news articles. Other aspects of health-related press releases persist as the main factors predicting variance in exaggeration. Undergraduate psychology students, so widely utilised for psychology research, may be perfectly adequate homologues for testing journalistic behaviour in experimental psychology settings.

CHAPTER 7 – GENERAL DISCUSSION

7.1. Summary

This thesis was concerned with the exaggeration of claims in the science media process. The impetus for this investigation began with discourse triggered by the coincidental link in the media between the 2011 British riots and a research paper reporting on localised neurotransmitter levels and impulsivity. From this discourse came forth the research (Sumner *et al.*, 2014) on which this thesis is based. The research demonstrated the relationship between press releases and news articles with regards to exaggerations of features that had relevance for human health-related behaviour. In order to support and expand on the findings of Sumner *et al.* (2014), I have utilised a number of different approaches, including: observational quantitative content analysis, interrupted time series analysis, a randomised controlled trial, experiments with specialist participants, qualitative thematic analysis, and online experiments. Using a number of different approaches has enabled the research to be seen from a number of different perspectives, but importantly it has also allowed me to develop myself significantly as a researcher.

Firstly I demonstrated that the results of Sumner *et al.* (2014) could be replicated with more recent data. The relationship between exaggeration in press releases and news was evident in my replication, as well as in the Dutch replication by Schat, Bossema, Numans, Smeets, and Burger (2018). In chapter three, an interrupted time series design was employed to test for a difference in article exaggeration before versus after the release of Sumner *et al.* (2014). The level of

exaggeration to statements of relationship in press releases after the publication of Sumner *et al.* (2014) were lower, suggesting that press officers had successfully become more cautious when describing the findings of observational research. Chapter four described a randomised controlled trial which aimed to directly test manipulations to the relationships and caveats in real world press release to attempt to track the effect on news articles. There was a great deal of condition non-adherence due to inclusion, by the press offices, of information we would have added to the press releases as part of the experimental manipulation. This meant that the only meaningful comparison to make was an as-treated observational analysis, which again strongly approximated the findings of Sumner *et al.* (2014). This meant that the experimental studies in chapters five and six were the remaining opportunities to test for a causal relationship between press release and news content. Chapter five described an experiment with journalism students that found that manipulations to statements of relationship and inclusion of caveats to study design did not have an effect on whether stories were selected for news. There was also no effect of statement of relationship manipulation on the statements of relationship written by participants. However, it was found that some press release items were more likely to be selected for news than others, meaning that factors other than the manipulations had an effect. Finally there was no effect of the type of writing objective given to participants of the proportion of exaggeration in statements of relationship.

7.2. Interpretation

Taken together, the observational findings from chapters two, three and four provide a support for the idea that the press releases are an important focus of the study of science communication. Press releases represent the first step in the ‘translation’ of complex scientific findings to the public. The uptake of information from press releases, but the lack of a relationship, in the UK data, between press release content and the generation of news suggests that UK press release should not be viewed as an opportunity to overstate claims in order to generate news. There is no benefit in a press release with an overstated claim, and no press release is better than an inaccurate press release. This supports the findings of Schwartz, Woloshin, Andrews, and Stukel (2012) that findings reported in accurate press releases are infrequently presented inaccurately in the news. Findings not reported by a press release can still be picked up from other sources and exaggerated in the news, but to a lesser degree than when a press release contains an inaccurate claim.

This suggests that the recommendation to press offices should be to follow guidelines on best practices in accurate reporting outlined by the Science Media Centre (2018), and the Academy of Medical Sciences (2018). This may seem somewhat like a recommendation to return to the deficit model, which pointed at a need for disseminating higher volumes of accurate science information to the public as a solution to waning public knowledge and interest. Dialogue with the public, and encouraging their participation in science are the dominant approaches, but it may be that the press release is too restrictive a medium to be used in a way that is conducive to the dialogue/participation models, since press

releases are not widely viewed by the public. The news however has mass public attention, and can use tools such as multimedia, comment-sections, social media, and advertising of events to encourage dialogue and participation. In this regard, if information transfer from press releases to news exists, then in addition to accurate reporting, press releases should be multimedia content rich, should advertise and coincide with events organised for the public, and should signpost opportunities to engage in dialogue with scientists about the research by, for example, including social media account information. The press release therefore should be a tool to assist with modern dialogue and participation approaches, but should ensure that focus is not drawn away from fundamental accuracy in reporting.

The lack of an experimental effect of press release content on news content suggests that other factors are more heavily involved in news synthesis than expected. The benefit of the experimental studies is that they cut out the majority of extraneous variables present in the real-world science communication process. It was thought that by removing these variables from the interaction between the press release and the news, that the experiments in this thesis would reveal the influence of press release exaggeration on news exaggeration. Person-to-person interactions in the science communication process (figure 1.1), independent investigation by the journalist, the journalist's own understanding and experiences, and the influence of their newsroom's perspectives are all potential influences of the story. This makes recommendations for future experimental research difficult. Our attempt at a randomised controlled trial in the real-world science communication process included these factors but failed because of the

difficulty of ensuring adherence to conditions by press offices; a problem that does not occur in my more controlled experiments of the simulated science communication process. It is unlikely that a study could be designed to avoid the issues present in both of these experimental approaches.

A positive outcome of these findings is that there appears to be no evidence of the churnalism behaviour in journalists suggested by Davies (2009). Masters journalism students did not directly reproduce the information manipulated in press releases. Rather, they produced restrained content: a low rate of exaggeration (around 30% of articles exaggerated) regardless of the exaggeration present in the source. This may be that the sample were acting in a way that was comparable to professional journalists but were not exposed to the other real-world influences that may be responsible for exaggeration, or alternatively the masters students are not representative of professional journalists because they were influenced by their academic knowledge and experience. In other words, they may be demonstrating textbook behaviours, literally. It is likely that their academic teachings focus on the avoidance of overstatement.

Data from UK studies showed that the uptake of stories into the news was not related to the presence of exaggeration. The attempted replication of this finding in the Netherlands showed that exaggerated articles were more likely to be picked up by the news. This is an interesting phenomenon, and an opportunity for further research. There are a number of possibilities for this finding. Firstly, it raises the question of whether differences in language might influence the perception of newsworthiness of exaggerated findings. My experimental finding

that differences in press releases other than the manipulations modulated their uptake, and that participants mainly referred to the ‘interestingness’ of a story, suggests that exaggerated language may be more interesting in Dutch. However, there should be no difference in language between exaggerated statements, and other non-exaggerated causal statements unless some causal statements are more interesting than others, and there was a difference in the frequency of interesting statements between exaggerated and non-exaggerated causal claims. In other words, would the word ‘impacts’ have more potential for news uptake than the word ‘ameliorates’? It is potentially an unlikely suggestion, but it would be relatively easy to retrospectively check the frequency of occurrence of words between conditions, or include this as a consideration in future studies. A second possibility for the difference between uptake across country may be the difference in sample. The majority of press releases in the Dutch sample came from university medical center press offices rather than the university press offices in the UK samples. This potentially means that the samples differ in the type of studies included, which could be a source of bias when interpreting the uptake potential of studies.

An opportunity for future research would be to test Dutch and UK journalists, or journalism students, with an amended method of the experimental study in chapter five. Formulaic press releases could be created, as outlined in chapter 5 rather than adapted from real-world press releases, such that other sources of variation in content could be controlled. This would also assist with matching press releases between Dutch and English. If the difference in news selection across languages still exists, then this may suggest that the source of the

difference in uptake is some aspect of language, rather than the source of the press release, or other unmeasured factors.

A lack of a difference in uptake across exaggeration in the UK studies may be indicative of a failure of the deficit model. More exaggerated press releases did not lead to more news uptake, which could mean that more exaggerated stories are not more appealing. But more accurate news stories also did not lead to more news uptake. This suggests that accuracy in reporting, the cornerstone of the deficit model, might not be more attractive to journalists for publication to the reader. Indeed, accuracy was not mentioned by journalism students in chapter five as a reason for or against selecting a news story. If accurate reporting was so important for science communication, and if accurate reporting is all that is needed for a better relationship between the public and science, then it would be expected that non-exaggerated findings would make it into news, to give the public what they want. Given that this was not found to be the case, and that the uptake of stories by journalists may be based on other factors, it would be useful to study whether the stories that do make it into the news have a higher proportion of content related to the dialogue model of science communication. In other words, are the stories that make it to news the ones that involve the public, or the ones that encourage dialogue and debate? A way to test this might be to perform a qualitative analysis of the content of press releases that do and do not have news uptake to see whether there are themes present in articles with uptake that are not present in those without.

7.3. Generalisation

Health-related news was initially selected for this research because of its potential impact on the public, should they be influenced by exaggeration in news stories. To understand whether this research would be generalisable to other forms of media, other news topics, and other aspects of science communication, we should consider the similarities and differences between health-related news and those mediums. This research investigated health news in the form of online and print news, but the randomised controlled trial also included television broadcast transcriptions. No comparison was made between the content of news in text versus broadcast, but anecdotally, I experienced a great degree of similarity between the articles in print and broadcast. The inverted pyramid style of writing, with the conclusions presented first followed by information of decreasing importance, was still present in the broadcast media, suggesting that print and broadcast health-news share similarities. With television programmes representing the number one source of science news for the public (Castell *et al.*, 2014) this further highlights the importance of press release accuracy.

The press releases used in this thesis were all written and published by university press offices, but the relationship between press offices and news has also been found using press releases from scientific journals (Sumner *et al.*, 2016). Though it would be expected that press releases from pharmaceutical companies would be similar, they actually infrequently report on published studies, and frequently only report basic details, so they may differ in their potential influence on news content (Kuriya, Schneid, & Bell, 2008).

The extent to which this research is relevant to other important news topics such as finance and politics is not clear, given the differences in content between topics. This research heavily focuses on the statement of relationship between variables as a measure of comparison between articles. Statements of relationship are in essence statements of certainty about the relatedness of variables. Political stories by comparison generally report the differing political perspectives addressing an issue, or highlight and critique the policies of people and groups. In comparison to health-news, there are essentially very few truths in political stories that can be represented by statements of relationship in the same way as scientific findings. Political news articles are also far less reliant on public relations material than health news. Around 40% of health and nature-related news articles were found to be mainly reliant on press relations material, in comparison to less than 10% of political news stories (Lewis, Williams, & Franklin, 2008).

7.4. Observational research

A large proportion of my time working on this project has been devoted to the coding and analysis of news articles for the two large datasets generated for the replication/interrupted time series analysis (data from 2014 to 2015), and the randomised controlled trial (data from 2016 to 2017). During this time I have understood observational data from two unique perspectives. In addition to being involved in large scale, well-designed observational studies, I also read and coded news reports of observational findings almost every day. As such I have been exposed to a vast quantity of questionable reports. For instance, one front-

page headline during the last general election claimed that poor diet was more lethal than smoking (Donnelly, 2015). In other words, eating junk food kills more people than does smoking. A striking result, if it were true. The report was actually based on an NHS report stating that poor diet had been listed in 10.8% of hospital admissions as a contributory factor to the illness of the patient, whereas smoking had been listed in only 10.7% of cases (Newton, 2015). Fault can be found with essentially every aspect of this report, not least the implicit suggestion equating hospital admissions with death. I presented this headline as part of an internal seminar and was asked the question “do you think it would be justified to only issue press releases for causal studies?” This could indeed be helpful, given the rates of exaggeration seen in news for exaggerated press releases reporting on observational research. The proposal from the questioner would be an abrupt way to potentially remove a good deal of exaggerated information from the science-media process. Though this suggestion could be inappropriate given the strength of evidence provided by some observational studies.

There are a great deal of well controlled, powerful, observational studies, that can often provide as compelling data as any experiment (Rubin, 2007). Indeed, reporting on the findings of the original research paper (Sumner *et al.*, 2014) and my replication and interrupted time series analysis, I discovered first hand how difficult it is to avoid inadvertently stating a causal claim. Such large-scale observational studies are often so compelling. Without some overly strong inferences made in response to the findings of Doll and Hill (1956), that smoking was associated with lung cancer and early death, we might have not made as

speedy progress with regards to our understanding of the problems with smoking as we have. A randomised controlled trial investigating the mortality of smokers would be extremely difficult because of ethical issues assigning cessation and smoking directions, the required length of the study, and condition non-adherence.

Arguments could also be made regarding the overall reduction in science coverage likely if only valid causal conclusions were communicated.

Observational research represents a large quantity of the findings that make it into the news. This could be seen as a negative; an overall reduction would mean that the public have less opportunity for exposure to science. On the other hand, the findings that do make it into the news would have a greater level of accuracy. The deficit model would be split by this argument; if a high volume of accurate information is recommended, it is unclear whether a higher volume of news with lower accuracy, or a lower volume of news with higher accuracy is most beneficial.

The considerations so far have only been how to approach science communication that is based on a single source. Schwitzer (2014) suggested that focus should be given to eliminating single source communication in favour of independent vetting of scientific claims using multiple sources, in addition to focussing on clinical end-points rather than surrogate markers. In other words reporting about mortality or elimination of pathology, rather than reporting about the reduction of a hormones related to a disease for example. Other suggestions

include reducing the use of anecdote as a source of evidence rather than illustration, and stating absolute risk rather than relative risk Schwitzer (2014).

7.5. Concordat on accuracy

The remedy to exaggeration in the media is of course not censorship of scientific reports. There are some fields of enquiry that rely almost solely on the compound effect of multiple strong observational research findings over time, such as respiratory health. The approach should be what has been advocated throughout this thesis. Health-related research should be reported accurately in an attempt to avoid the possibility that exaggeration can originate in the press release; it is not certain how exaggeration in press releases makes it into the news, but press releases should not give news the opportunity to absorb inaccuracy. As I reported with regard to the Concordat on Openness on Animal Research in the UK (2014), press offices have a unique opportunity in that they have a now very strong support network in the form of the Academy of Medical Sciences, The Science Media Centre, and STEMPRA. Just as a number of organisations came together to commit to more transparent and accurate reporting of information relating to animal research, the infrastructure is in place for a similar movement relating to the reduction of exaggeration in the media. The issues of condition non-adherence in the randomised controlled trial in chapter four suggest that press offices are potentially already incorporating the suggestions from the Academy of Medical Sciences and STEMPRA in their work. They have started to include more qualification of study design by way of mentioning caveats, and they have reduced exaggeration in statements of relationship. This positive change was not complete, there was still exaggeration and a lack of caveats in

press releases, but a potential barrier to more complete improvement might be the high turnover of staff members in press offices. During the randomised controlled trial it was noted by my colleagues that the point of contact for our communication with each press office during the trial changed frequently such that the justification for the study had to be reiterated frequently. This could mean that the improvements in those press offices that have been exposed to the guidelines by STEMPRA and the Academy of Medical Sciences may be transient.

This thesis did not manage to uncover the nuances of the transmission of exaggeration through the science media process to the extent that it set out to, but it did confirm that there is no penalty for a reduction of exaggeration in health-related communication. Furthermore, the selection of news stories appears to be more likely related to the newsroom's appetite for other aspects of health-stories that might be beyond the press offices control. If newsrooms are going to cover a particular story, it is the responsibility of the press office to ensure that the material available to the journalist does not give them the opportunity to exaggerate. The involvement of scientists should not be underestimated, and scientists should take care regarding their involvement at every step of the science communication process. This research indicated a high level of exaggeration even in the journal article, so scientists should consider that their involvement in the science communication process starts by ensuring that the language in their studies is appropriate for the study design. At the level of the press release, it would be useful for scientists to acknowledge their responsibility to communicate with press officers and to be aware of the STEMPRA and Academy of Medical Sciences guidelines in order for them to understand what

information to provide to press officers to help them avoid exaggeration at the level of the press release. Scientists should embrace discourse at the level of the news article by participating in dialogue with journalists to provide a further level of vetting of exaggeration of their research. And finally, scientists should understand that the communication of science is more involved than merely the dissemination of accurate findings. Dialogue with the public, and their participation in outreach events represent a key opportunity for impact.

7.6. Conclusion

Exaggeration in health-news is linked to exaggeration in press releases, but is possibly mediated by a range of other factors in the science communication process. The pressures of the need for timely publication, impact, and research funding for scientists; the need for the fast turnaround of engaging content for raising institutional profile for press officers; and the need for the fast turnaround of a high volume of news coverage for journalists, mean that there are many opportunities for inaccuracy to enter the science communication process. It appears that the opportunities for improvement of science communication are a few small changes to the practice of scientists and press officers. The suggestions of STEMPRA and the Academy of Medical Sciences regarding the accuracy in statements of relationship and the inclusion of details regarding study design are simple to follow, and may help to increase the accuracy of science communication.

This should be seen as an empowering message to science communicators.

References

- Academy of Medical Sciences. (2018). AMS press release labelling system for new medical research. Retrieved from <http://www.sciencemediacentre.org/wp-content/uploads/2018/01/AMS-press-release-labelling-system-GUIDANCE.pdf>
- Adams, R. C., Challenger, A., Bratton, L., Boivin, J., Bott, J., Powell, G., Williams, A., Chambers, C., & Sumner, P. (2019). Claims of causality in health news: a randomised trial. *BMC Medicine*, *17*, 91.
- Adams, R. C., Sumner, P., Vivian-Griffiths, S., Barrington, A., Williams, A., Boivin, J., Chambers, C. D., Bott, J. (2017). How readers understand causal and correlational expressions used in news headlines. *Journal of Experimental Psychology: Applied*, *Vol 23*(1), 1-14.
- Adelman, R. C., & Verbrugge, L. M. (2000). Death makes news: The social impact of disease on newspaper coverage. *Journal of Health and Social Behavior*, *41*(3), 347–367.
- Altmetric.com. (2018). <https://bmj.altmetric.com/details/2969665>
- Autism Network for Dietary Intervention News. (1998). Retrieved from <https://briandeer.com/wakefield/andi-interview.htm>
- Bjørnsen, G., Hovden, J. F., & Ottosen, R. (2007). Journalists in the making. *Journalism Practice*, *1*(3), 383-403.
- Bott, L., Bratton, L., Diaconu, B., Adams, R. C., Challenger, A., Boivin, J., . . . Sumner, P. (2019). Caveats in science-based news stories communicate caution without lowering interest. *Journal of Experimental Psychology: Applied*. Advance online publication.
- Boy, F., Evans, C. J., Edden, R. A. E., Lawrence, A. D., Singh, K. D., Husain, M., & Sumner, P. (2011). Dorso-lateral prefrontal γ -amino butyric acid in men predicts individual differences in rash impulsivity. *Biological Psychiatry*, *70*(9), 866–872.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.
- Brechman, J., Lee, C. J., & Cappella, J. N. (2009). Lost in translation?: A comparison of cancer-genetics reporting in the press release and its subsequent coverage in the press. *Science Communication*. *30*, 453–474.
- Brown, S. (2009). The New Deficit Model. *Nature Nanotechnology*, *4*, 609-611.

- Bultitude, K. (2010). Presenting Science. In Brake, M.L. and Weitkamp, E. (Eds) *Introducing Science Communication*. London: Palgrave MacMillan.
- Burwell, H., Canning, S., Dawson, R., et al. (2017). Stempra press officer guidelines 2017. Retrieved from https://stempra.org.uk/wp-content/themes/stempra/downloads/2017_stempra_guide_to_being_a_media_officer.pdf
- Castell, S., Charlton, A., Clemence, M., Pettigrew, N., Pope, S., Quigley, A., Jayesh Navin Shah, J. N., & Silman, T. (2014). Public attitudes to science 2014. Ipsos MORI.
- Davies, N. (2009). *Flat Earth News*. London: Vintage Books.
- Doll, R., & Hill, A. B. (1956). Lung cancer and other causes of death in relation to smoking: A second report on the mortality of British doctors. *British Medical Journal*. 2(5001), 1071–1081.
- Domenighetti, G., Luraschi, P., Gutzwiller, F., Pedrinis, E., Casabianca, A., Spinelli, A., & Repetto, F. (1988). Effect of information campaign by the mass media on hysterectomy rates. *Lancet*. 332 (8626-8627), 1470-1473.
- Donnelly, L. (2015, September 15). Poor Diet More Lethal Than Smoking. *The Daily Telegraph*, p. 1.
- Donsbach, W. (2004). Psychology of news decisions: Factors behind journalists' professional behavior. *Journalism*. 5(2) 131–157.
- Einsiedel, E. (2007). Editorial: Of publics and science. *Public Understanding of Science*, 16, 5-6.
- Entwistle, V. (1995). Reporting research in medical journals and newspapers. *BMJ*. 310, 920-923.
- Felt, U. (2003). Why Should the Public 'Understand' Science? A Historical Perspective on Aspects of the Public Understanding of Science, in M. Dierkens and C. Von Grote (eds) *Between Understanding and Trust: The Public, Science and Technology*, reprinted 2003, 7–39. London: Routledge.
- Fox, S. (2006, October 29). Online Health Search 2006. *PEW Internet & American Life project*.
- Fox, S., & Rainie, L. (2002). Vital decisions: How Internet users decide what information to trust when they or their loved ones are sick. *Pew Internet & American Life Project*.

- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on “Estimating the reproducibility of psychological science”. *Science*, *351*, 1037.
- Godlee, F., Smith, J., & Marcovitch, H. (2011). Wakefield’s article linking MMR vaccine and autism was fraudulent. *BMJ*, *342*, c7452
- Grilli, R., Ramsay, C., & Minozzi, S. (2002). Mass media interventions: effects on health services utilisation. *Cochrane Database Systematic Review*, *1*:CD000389
- Health news coverage in the US media. (2008). A report by the Kaiser Family Foundation and the Pew Research Centre’s Project for Excellence in Journalism.
- Hyland, K. (1996). Talking to the academy: Forms of hedging in science research articles. *Written Communication*, *13*, 251-281.
- Ireland, M. E., & Pennebaker, J. W. (2010). Language Style Matching in Writing: Synchrony in Essays, Correspondence, and Poetry. *Journal of Personality and Social Psychology*, *99*(3) 549–571.
- IRI. (2015, November 23). Processed meat cancer scare wipes 10% (£3 million) off sales of sausages and bacon according to IRI Retail Advantage data press release about reduction in meat sales after WHO report. Retrieved from [https://www.iriworldwide.com/en-GB/insights/news/Processed-meat-cancer-scare-wipes-10-\(£3-million\)](https://www.iriworldwide.com/en-GB/insights/news/Processed-meat-cancer-scare-wipes-10-(£3-million))
- Keohane, R. O., Lane, M., & Oppenheimer, M. (2014). The ethics of scientific communication under uncertainty. *Politics, Philosophy, & Economics*, *13*(4).
- Kuriya, B., Schneid, E. C., & Bell, C. M. (2008). Quality of pharmaceutical industry press releases based on original research. *PLoS ONE*, *3*(7).
- Lewis, J., Williams, A., & Franklin, B. (2008a). Four rumours and an explanation. *Journalism Practice*, *2*(1), 27-45.
- Lewis, J., Williams, A., & Franklin, B. (2008b). A compromised fourth estate? *Journalism Studies*, *9*(1), 1-20.
- Li, M., Chapman, S., Agho, K., & Eastman, C. J. (2008). Can even minimal news coverage influence consumer health-related behaviour? A case study of iodized salt sales, Australia. *Health Education Research*, *23*(3), 543-548.
- Patterson, T. E., & Donsbach, W. (1996). News Decisions: Journalists as Partisan Actors. *Political Communication*, *13*, 455–68.
- Poland G. A., & Spier, R. (2010). Fear, misinformation, and innumerates: How the Wakefield paper, the press, and advocacy groups damaged the public health. *Vaccine*, *28*, 2361–2362.

- Maclure, M., Dormuth, C., Naumann, T., McCormack, J., Rangno, R., Whiteside, C., & Wright, J. M. (1998). Influences of educational interventions and adverse news about calcium-channel blockers on first-line prescribing of antihypertensive drugs to elderly people in British Columbia. *Lancet* 352(9132), 943–948.
- McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC Medical Research Methodology*, 7, 30.
- McVey, D., & Stapleton, J. (2000). Can anti-smoking television advertising affect smoking behaviour? Controlled trial of the Health Education Authority for England's anti-smoking TV campaign. *Tobacco Control* 9, 273–282.
- Medvecky, F., & Leach, J. (2017). The ethics of science communication. *Journal of Science Communication*, 16(4).
- Nattinger, A. B., Hoffmann, R. G., Howell-Pelz, A., & Goodwin, J. S. (1998). Effect of Nancy Reagan's mastectomy on choice of surgery for breast cancer by US women. *JAMA* 279(10), 762–766
- Nelkin, D. (1995). *Selling science*. New York. W. H. Freeman.
- Newton, J. N., Briggs, A. D. M., Murray, C. J. L., Dicker, D., Foreman, K. J., Wang, H. D., . . . Davies, A. C. J. (2015). Changes in health in England, with analysis by English regions and areas of deprivation, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386(10010), 2257-2274.
- Norris, S., Phillips, L., & Korpan, C. (2003). University students' interpretation of media reports of science and its relationship to background knowledge, interest, and reading difficulty. *Public Understanding of Science*, 12, 123–145.
- Parker, A. A., Staggs, W., & Dayan, G. H. (2005). Implications of a 2005 Measles Outbreak in Indiana for Sustained Elimination of Measles in the United States. *The New England Journal of Medicine*. 355(5), 447-455.
- Patterson, T. E., & Donsbach, W. (1996). News Decisions: Journalists as Partisan Actors, *Political Communication* 13, 455–68.
- Pottker, H. (2003). News and its communicative quality: the inverted pyramid—when and why did it appear? *Journalism Studies*, 4, 501-511.
- Research Councils UK (2010). *What's in it for me? The benefits of public engagement for researchers*. Research Councils UK. Retrieved from <https://webarchive.nationalarchives.gov.uk/20101112204105/http://www.rcuk.ac.uk/per/bestprac.htm>

- Rogers, S., Sedghi, A., & Evans, L. (2011). UK riots: every verified incident - interactive map. The Guardian. Retrieved from <https://www.theguardian.com/news/datablog/interactive/2011/aug/09/uk-riots-incident-map>
- The Royal Institution. (2012, March 30). Alok Jha: Science and the Media - Presentations (1/2) [Video file]. Retrieved from https://www.youtube.com/watch?v=v8uDnSm_W9U
- Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine*, 26, 20-36.
- The Russell Group. (2014, December 18). Research Excellence Framework. Retrieved from: <https://russellgroup.ac.uk/news/research-excellence-framework/>
- Schat, J., Bossema, F. G., Numans, M. E., Smeets, I., & Burger, J. P. (2018) Overdreven gezondheidsnieuws. Relatie tussen overdrijving in academische persberichten en in nieuwsmedia, *Nederlands Tijdschrift voor Geneeskunde* 162, 13-17.
- Schwartz, L. M., Woloshin, S., Andrews, A., & Stukel, T. A. (2012). Influence of medical journal press releases on the quality of associated newspaper coverage: retrospective cohort study. *BMJ*, 344.
- Schwitzer, G. (2015). Trying to drink from a fire hose: Too much of the wrong kind of health care news. *Trends in Pharmacological Sciences*. 36, 623-627.
- Science Media Centre. (2018). SMC best practice guidelines. Retrieved from <http://www.sciencemediacentre.org/wp-content/uploads/2012/09/10-best-practice-guidelines-for-science-and-health-reporting.pdf>
- Scott, J. (2012, March 14). Scientists and journalists need different things from science... or do they? London Blog. Retrieved from <http://blogs.nature.com/london/2012/03/14/scientists-and-journalists-need-different-things-from-science-or-do-they>
- Selk, A. (2017, November, 18). Please stop annoying this NASA scientist with your ridiculous Planet X doomsday theories. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/speaking-of-science/wp/2017/11/18/please-stop-annoying-this-nasa-scientist-with-your-ridiculous-planet-x-doomsday-theories/?noredirect=on&utm_term=.035699d97e8c

- Sibbald, B., & Roland, M. (1998) Understanding controlled trials: Why are randomised controlled trials important? *BMJ*, 316, 201.
- Slapak, M. (2004). The effect of The World Transplant Games on transplant rates in five continents. *Annals of Transplantation*. 9, 46–50.
- Society of Professional Journalists (2014). *SPJ Code of Ethics*. Retrieved from <http://www.spj.org/ethicscode.asp>
- Soumerai, S. B., Ross-Degnan, D., & Kahn, J. S. (1992). Effects of professional and media warnings about the association between aspirin use in children and Reye's syndrome. *Milbank Quarterly* 70(1), 155–82
- Starcevic, V., & Berle, D. (2013). Cyberchondria: towards a better understanding of excessive health-related Internet use. *Expert Review of Neurotherapeutics*, 13(2).
- Starck, K., & Soloski, J. (1977). Effect of Reporter Predisposition in Covering Controversial Stories, *Journalism Quarterly* 54, 120-125.
- Sturgis, P., & Allum, N. (2004). Science in society: re-evaluating the deficit model of public attitudes. *Public Understanding of Science*. 13, 55–74.
- Sugerman, D. E., Barskey, A. E., Delea, M. G., et al. (2008) Measles outbreak in a highly vaccinated population, San Diego, 2008: role of the intentionally undervaccinated. *Pediatrics*. 125(4), 747–755.
- Sumner, P., Boy, F., & Chambers, D. C. (2011, August 22). Riot control: How can we stop newspapers distorting science? *The Guardian*. Retrieved from <https://www.theguardian.com/science/blog/2011/aug/22/riot-control-newspapers-distorting-science>
- Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Bott, L., Adams, R., ... Chambers, C. D. (2016). Exaggerations and Caveats in Press Releases and Health-Related Science News. *PLoS One*, 11(12).
- Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C. A., Davies, A., ... Chambers, C. D. (2014). The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ*, 349.
- Tanner, A. H. (2004). Agenda building, source selection, and health news at local television stations. *Science Communication*, 25(4), 350-363
- Taylor, J. W., Long, M., Ashley, E., Denning, A., Gout, B., Hansen, K., ... Newton, P. M. (2015). When Medical News Comes from Press Releases—A Case Study of Pancreatic Cancer and Processed Meat. *PLoS ONE*, 10(6), e0127848.

- Tesoriero, J. M., & Sorin, M. D. (1992). The effect of 'Magic' Johnson's HIV disclosure on anonymous HIV counselling and testing services in New York State. *AIDS & Public Policy Journal* 7, 216–24.
- Van der Sanden, M. C. A., & Meijman, F. J. (2008). Dialogue guides awareness and understanding of science: an essay on different goals of dialogue leading to different science communication approaches. *Public Understanding of Science*, 17, 89-103.
- Van der Worp, H. B., Howells, D. W., Sena, E.S., Porritt, M.J., Rewell, S., O'Collins, V., et al. (2010). Can animal models of disease reliably inform human studies? *PLoS Med* 7, e1000245.
- Wakefield, M. A., Loken, B., & Hornick, R. C. (2010). Use of mass media campaigns to change health behaviour. *Lancet*, 376, 1261–1271.
- Wilkins, L. (2006). Plagues, Pestilence and Pathogens: The Ethical Implications of News Reporting of a World Health Crisis. *Asian Journal of Communication*, 15(3).
- Williams, A., & Clifford, S. (2010). Mapping the field: Specialist news journalism in the UK national Media.
- Willoughby, T., Adachi, P. J. C., & Good, M. (2012). A longitudinal study of the association between violent video game play and aggression among adolescents. *Developmental Psychology*, 48(4), 1044-1057.
- Wilson, A., Bonevski, B., Jones, A., & Henry, D. (2009). Media reporting of health interventions: Signs of improvement, but major problems persist. *PloS ONE*. 4(3).
- Young, P. (2017). Research Access and Discovery in University News Releases: A Case Study. *Journal of Librarianship and Scholarly Communication*, 5(1).
- Zach, L., Dalrymple, P. W., Rogers, M. L., & Williver-Farr, H. (2012). Assessing Internet access and use in a medically underserved population: implications for providing enhanced health information services. *Health Information and Libraries Journal*, 29(1), 61-71.
- Zickuhr, L. (2010). Generations 2010. *Pew Internet and American Life Project*.

Appendices

Appendix 2.1.

Coding sheet devised by the InSciOut research group to code journal articles, press releases and news articles.

	A	B	C	D	E	F	G	H
1	Press Release Title							
2	Study Title: Original Research							
3	Discipline							
4	Sub-discipline							
5	Sub-discipline 2							
6	Notes (complete/incomplete)							
7	Coder:			Draft Press Release	Final Press Release	Journal Article - Title and Abstract	Journal Article - Body	Daily Mail
8	Reference:		Available for Coding?					
9			Date					
10			Who is news article written by?					
11			What's the actual job title of the journalist?					
12		Press Release Characteristics	Expected Level of News Uptake					
13			Particular Desired Outcome					
14			Joint Release					
15			Embargo Start Day					
16			Embargo Time					
17			Embargo Duration (in days)					

	A	B	C	D	E	F	G
18		IV/DV	Main conclusion IV(s)				
19			Main conclusions DV(s)				
20			Are the IV and DV the same (between JA and PR and between JA and NA)				
21		TITLE	Primary statement of relationship				
22		Causation/Relationship about main IV and DV	Statement of Relationship Code				
23		MAIN STATEMENTS	Primary statement of relationship				
24		Causation/Relationship about main IV and DV	Statement of Relationship Code				
25			Are the IV and DV the same in the title and the main statements?				
26		Sample	ACTUAL SAMPLE				
27			Who are the Main Conclusions About?				
28			Sample Code				
29		Study Design info	Study Design				
30			Any information about study design?				
31			Study design stated?				
32			Mention whether cause and effect can be deduced?				
33			If they mention that cause and effect cannot be deduced - do they state why?				
34			Study context mentioned				
35			Any indication of the strength/ weakness of a study?				
36		RCT only	Statement of study design				
37			words to mention				
38			included in a quote?				
39		Advice	Advice to Change Behaviour				
40			Code				
41		Other	Press offices' institution mentioned?				
42			words before mention				
43			body word count (exc. Title)				
44		PR Manipulations	Experimental condition				
45			Synonym change				
46			words to mention				
47			Causal phrasing change title				
48			words to mention				
49			Causal phrasing change main statements 1.				
50			words to mention				
51			Causal phrasing change main statements 2.				
52			words to mention				
53			Design statement inserted (text)				
54			included in a quote?				
55			words to mention				

Appendix 5.1

Coding sheet I adapted for the purpose of coding participant data in the experimental studies. This is adapted from that created by the InSciOut team.

A	B	C	D	E
Press Release	Antidepressants			
Coder:	LB		Final Press Release	example
	Word count	Title word count		12
		Body word count		120
	IV/DV	Does participant report same variables as PR		1
		Main conclusion IV(s) TEXT	antidepressants	pills
		Main conclusions DV(s) TEXT	severe side effects	death
	TITLE	Statement of relationship TEXT		Depression pills lead to untimely deaths
	Causation/Relationship about main IV and DV	Statement of Relationship CODE		6
	MAIN STATEMENTS	Statement of relationship TEXT		the report suggests that deaths have been linked to the new-style pills
	Causation/Relationship about main IV and DV	Statement of Relationship CODE		3

Appendix 5.2

Participants raw free text comments regarding in response to the question: ‘Did you use/find useful any sources other than those provided? If so, which? (e.g. other news websites, journal webpages etc.)’, and in response to the question: ‘To what extent did you use/find useful the original journal articles?’.

Did you use/find useful any sources other than those provided? If so, which? (e.g. other news websites, journal webpages etc.)	Reason/which
I did for task 2. http://www.who.int/cardiovascular_diseases/en/ http://www.livescience.com/51622-teens-emergency-contraception-increase.html https://www.theguardian.com/business/2017/mar/07/nestle-sugar-remove-uk-ireland-2018	Other news. Scientific body webpage.
Task 1) Again, I wasn’t able to make use of them in the time period provided. Task 2) Again, I wasn’t able to make use of them in the time period provided.	Time constraint
No, as that would have manipulated the results, which would essentially render them invalid.	Comment on experiment?
I didn’t look anywhere else, partly because of the time constraints. Also usually a press release and the report is enough anyway. I sometimes look at how the BBC words articles, if I had more time I may have looked at other news organization news outlets.	Time constraint. Other news. Press release is sufficient.
Didn’t use any other sources – press releases only.	-
I didn’t use any other sources other than the press release.	-
I used the internet to check the official title of the Thorax publication, in order to better help me give it an appropriate and relatable title in my piece. I also checked the current school of thought surrounding the risks of developing osteoporosis linked to certain exercise, but little was available in the allotted time.	Journal page. Background reading.
News websites to double check if there were any previous cases related to some of the subjects and a dictionary, as some texts were too technical.	Other news. Define terms
I didn’t have the time to, but it did come to mind because I would need to ‘balance’ every article I write.	Time constraint
no	
I didn’t really have time to look for other sources. Tended to stick to just the press releases and journal articles.	Time constraint
I checked the acronym for Department of Public Health and Primary Care, that was it.	Define terms
Under the time constraints, I did not use them. All information was in the press release.	Time constraint. Press release is

	sufficient.
no	
I did not use any other sources apart from google to double check the definition of osteoporosis.	Define terms
no	

To what extent did you use/find useful the original journal articles?	Response
If I had more time I would have used them more. I did use them a lot in the first few questions of task 1 but this slowed my question-answering speed considerably. I used them to check that what the press release was suggesting was actually similar to the journal findings.	Time constraint. Fact checking
Task 1) In the time period available, I wasn't able to look at them at all. Task 2) Having read the press-releases once already, I endeavored to read the original articles this time around. Although there is danger in my lack of scientific background leading me to misinterpret the information, I nonetheless feel it behoves me to look – as the press releases alone may be misleading, in either overstating their findings or oversimplifying the data. I found it useful to see the original version, as expected, as it did give a clearer picture than the press release which, I found, gave baffling precedence to certain parts of the findings over others. No journalist should try to write a story without finding the original source. A good example of this was recently when a story about research into male contraceptive injections was widely misreported by the press as, instead of looking at the original study, they instead just re-wrote each other's stories – leading to the published results resembling the original article less and less in each retelling and causing a wide reaction online – based on wrong information!	Time constraint. Potential for misinterpretation.
Not very as they were too long to read in the given time frame.	Time constraint
The time constraints in task 1 especially meant I didn't look at most of them. I didn't have time to go through them as I would if I was writing an article. I would have liked to to	Time constraint. Fact checking.

ensure my answers were based on the facts. In any story that has research the report needs to be skim read for the important facts.	
Didn't use them – not enough time to read both the journal article, and the press release, so decided to focus on the press release when determining newsworthiness.	Time constraint
I didn't really use them, when you have a limited time to write the story or judge a press release on newsworthiness then the key information needs to be in the press release. I would use the original journal article if it was a part of a wider bigger story, but in this case the press release was all that was needed. In terms of the first task, judging if the story is newsworthy, then if it can't be easily summed up in a press release then it is unlikely to be a story worth covering in a short amount of time.	Key information should be in the press release; not newsworthy if more than the press release is needed. Time constraint.
The original articles had some use to clarify figures, but in terms of ingesting, processing and analyzing the original articles, time constraints meant much of my analysis came from the press releases themselves. Though the original articles were referred to briefly, they could have been omitted entirely and I would have still formulated much the same articles around them.	Figures useful. Time constraint. Assumption that the press release is a good representation.
There is some information I didn't know about and that it may be interesting to write about as well.	More information in journal article.
Useful to the point that it is their findings and they provided the information they wanted to get out, but some of them were slightly 'technical' I would have wanted to get in touch with them to find out more so as to help me simplify my task of informing the public better.	Too technical. Need to contact scientists (constraint of experiment)
In general I found them quite confusing. They were extremely hard to read and full of educational and professional language that made it hard to decipher for people like me who were unsure of the topic in question. Trying to pick out the important facts from these journal articles was different as there was so many numbers and important looking facts jumbled over the pages.	Journal article is confusing and specialised.
The journal articles were less useful than the press releases, the press releases tended to	Journal article too complicated. Press release is concise.

offer the information in a more succinct and user friendly way. Ultimately, scientific articles can be difficult to process for journalists and indeed more difficult to translate to readers. The press release mediate this information.	
At the start of the first exercise, I found them useful and interesting, but I found that I was using up too much time reading them, especially if I found them interesting. As time ran out, I stuck to just reading the press releases.	Journal article was useful, but time constraint.
There was not enough time to read beyond the abstract, at most. This is worrying, because it means that whoever writes the press release has far too much control. The more technical the language or more terms of art included in a source, the less time there actually is to assess it properly. This means that journalists are aware they are not in a good position, which means that they may make it more difficult for PRs out of insecurity. It's not a case of saying, "well, they're scientists aren't they? They wouldn't try and get in the paper without something important." These press releases are written by POs and that's the bottom line.	Time constraint. No choice but to use the PR.
Under the time constraints, I did not use them. All information was in the press release.	Time constraint. Confident in press release
Some were a little bit useful but probably would have been easier to understand if I had medical knowledge or was a health reporter. Was good to read aims of the study to get a feel for what they were trying to come up with.	Confusing
Only when looking at the results. For the vast majority I opened the journals briefly before relying on the press release. I do enjoy reading scientific journals and I'm sure that having the time available to read them to completion would have benefit the articles. While writing for news I'm sure that this is rarely possible for most publications.	Time constraint.
I found the press releases easier to understand and be able to write from.	Ease of understanding