THE UNIVERSITY of York

# How does clickbait work:

## An eye-tracking method to discover people's reactions

| | | |
|---|---|---|
| Student Name | : | Yiteng Xing |
| Date | : | 19th September 2016 |
| Supervisor | : | Dr. Alistair Edwards |
| Word count | : | 12312 |

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Human-Centred Interactive Technologies
in the Department of Computer Science
at the University of York

Number of words = 12312, as counted by the MS Word word count command. This includes all the body of the report, but excludes Appendix.

# ABSTRACT

Clickbait web content is designed to entice people into clicking on a link. The publishers use strategies in writing headlines in order to create a 'curiosity gap' between the content and the reader. However, this bait often tricks people into clicking and leaves them disappointed. As a growing number of people are getting information through online media platforms, it is necessary to discover how clickbait works and which elements induce people to click on it. In this paper, we designed a lab-based experiment to observe the participant's reaction to clickbait by using screen recording and eye-tracking mechanisms. A qualitative research method with two coding schemes was applied to 151 headlines which were clicked by 20 participants. The results show that people are attracted by clickbait headlines that contain elements of 'a list', 'game', 'celebrity' and 'You or I'. However, the result changes slightly based on the different types of user. Then, the clickbait working principle, effectiveness of the eye-tracking method and the limitations of this study are discussed. In addition, this study contributes to further work by generating sufficient categorisation for clickbait and more specific research can be conducted such as clickbait with culture, language and social network sharing.

# ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to all those who provided me with the chance to complete this report. I would especially like to thank my project supervisor, Dr. Alistair Edwards, who gave me very detailed suggestions every week over the past five months and helped me to coordinate my project, particularly in writing this report. I would also like to extend my deepest gratitude to all those who have directly and indirectly guided us in writing this assignment. Many people, especially some of our HCI group team members – Professor Helen Petrie, Mr. Leonardo Guzman and Mr. Andrew Lewis, have made valuable comments or suggestions on the experiment, which gave me the inspiration to improve this paper. I also appreciate the University of York for providing me with the eye tracker to complete the experiment. I sincerely thank everyone for their direct and indirect help in completing my final project.

# CONTENTS

ABSTRACT..... 2

ACKNOWLEDGEMENTS..... 3

LIST OF TABLES..... 5

LIST OF FIGURES..... 6

CHAPTER 1: Introduction..... 7

CHAPTER 2: Literature Review..... 12

CHAPTER 3: Methodology.....26

CHAPTER 4: Result..... 40

CHAPTER 5: Discussion..... 53

CHAPTER 6: Conclusion..... 66

REFERENCES.....68

APPENDIX A: Consent Form.....74

APPENDIX B: Information Sheet.....76

APPENDIX C: Debriefing Sheet.....79

# LIST OF TABLES

Table 1. First coding scheme of 151 headlines with images…….….. 37

Table 2. Second coding scheme of 151 headlines with images……...……37

Table 3. Overview of coding data with occurrence and percentage…….…... 41

Table 4. Percentage of different click preferences between male and female……..43

Table 5. Percentage of different click preferences between two age groups..….….45

Table 6. Overview of the second coding result……...…..…... 47

Table 7. Comparison of Non-clickbait and clickbait headlines based on number of coding elements per title…….…..48

Table 8. Overall comparison between Section 1 and Section 3……...….49

Table 9. Gender group click preference comparison between Section 1 and Section 3…….51

Table 10. Gender group click preference comparison between Section 1 and Section 3……...….52

# LIST OF FIGURES

Figure 1. News (a) and Clickbait (b) headline formats…….….8

Figure 2. A click through rate comparison between two types of advertisement…..... 25

Figure 3. Percentage of Headlines that contain these 19 Coding Elements……... 41

Figure 4. Percentage of different click preferences between male and female.……...43

Figure 5. Percentage of different click preferences between two age groups...…..….46

Figure 6. Overview of the second coding result……...…...….... 47

Figure 7. Comparison of Non-clickbait and clickbait headlines based on number of coding elements per title…….…..48

Figure 8. Overall comparison between Section 1 and Section 3……....….50

# CHAPTER 1

## Introduction

One of the most important considerations for online journalists is attracting a reader's attention through headlines alone. Without interest in a headline, there will be no interest in the story. By enticing a reader to click on a catchy headline, there is a high click rate for the website, which leads to advantages for the website owners.

There are various stances that journalists take in trying to grab a person's attention. Schaffer notes the use of gossip-like content that highlights aspects such as scandal, sex, heartbreak, or even maximising on interest of the supernatural [1]. Similarly, Molek-Kozakowska and Wallberg identify the role of sensationalism and provocation in these headlines. It does not purely come down to content [2] [3], however; Lindemann notes the use of narrative techniques to increase the appeal of a headline [4].

Ultimately, if a headline does not capture a reader's attention, then they will not be driven to read the entire article, so headlines must be particularly clever or alluring. This has always been a generally accepted fact, however the widespread use of the Internet in modern society has

created an even more important need to draw in a reader through catchy headlines, known as “clickbait”.

Figure 1 below demonstrates examples of ‘Clickbait’, described as a “sensationalised headline or piece of text on the Internet designed to entice people to follow a link to an article on another web page” [5]. Clickbait is often misrepresentative and therefore risks spreading distortion of truth. For example, Figure 1 compares two headlines written on the same news piece, the publication of Michelle Obama’s biography. It can clearly be seen how (b) has chosen to highlight the salacious aspects of the story whereas (a) takes a more academic approach.

**Figure 1. News (a) and Clickbait (b) headline formats [5]**

The widespread adoption of this style of journalism is already under criticism, that is, the seemingly sensationalised tabloid-style approach, and online clickbait is perceived to be even more dangerous [5]. Clickbait headlines condense certain aspects of a story, maximising manipulation of fact and using word-of-mouth as gospel. Although this approach clearly entices readers, it perpetuates gossip and mistruth, and journalistic integrity becomes ever weakened. The internet is a dangerous medium for news; stories are presented as an ensemble in "streams" so that reputable and non-reputable sources are mixed together and headlines are taken out of context [6]. Chen, Conroy and Rubin criticised Facebook for this presentation of news; online, news spreads quickly therefore rumours and warped truths can cause huge damage [5]. Consequently, Blom and Hansen call for clickbait to be analysed in terms of ethics rather than how much money it can generate [7]. Clickbait can be understood in terms of Chen, Conroy and Rubin's concept of "sense making metaphors" [5], which are used in the study of Library and Information Science to uncover how interest in a subject is at once created and fulfilled, as with clickbait [8]. Readers are given an intense sentiment of intrigue, not wanting to be disappointed in this quest for discovery and seek fulfilment for their desire for information by looking at more articles [9]. Clickbait headlines feed on this drive for knowledge, tempting readers to discover the full story through small, intriguing aspects. These features could be utilised by information systems

to detect lowbrow clickbait news. Indeed, former studies have attempted to introduce automatic systems to authenticate news, but have yet to do so for clickbait [5]. Chen, Conroy and Rubin discuss how connections are made between pieces of information to create a gap in understanding of a subject, and then fill that gap [5]. Sense making metaphors are essential considerations in understanding how intrigue is created by clickbait headlines and the lexical choices made.

However, getting more clicks and thereby more visits on websites ultimately amounts to one thing: these websites can raise their advertising rates and make profit [10]. Publishers often incentivise their writers to get clicks on stories, which can lead to the simplification of articles and headlines for monetary gain [10]. Slant, an online magazine, pays writers $5 for every 500 clicks on a story on top of their monthly salary (according to the Columbia Journalism Review) and this is a pattern followed by more and more websites [6] [10].

This analysis will uncover the function of 'clickbait' in practice and public response to such devices, by adopting a qualitative research method to record participants' online behaviour. This study is based on no previous research or theories, and its key consideration is the impact that factors such as age have on the effect of clickbait. Participants' online behaviour and eye-movement will be recorded using a software that can register this

information. One key issue to be held in consideration is the Hawthorne effect, and the way that this has been addressed will be explained further on in this essay.

# CHAPTER 2

# Literature Review

## 2.1. Clickbait: An Overview

According to Oxford dictionaries, Clickbait is defined as '(On the Internet) content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page' [11]. For example, "8 Things You Should NEVER Tell Your Work Friends", "Find Out Which Celebrity Doesn't Even Talk to Her Husband", "Have You Seen the 15 Biggest Homes in Hollywood?" and "You Won't Believe What This Supermarket Discovered"

Often associated with the online tabloid press, clickbait headlines generally focus on subjects such as celebrities, rumour and fictional accounts, rather than hard-hitting journalism on more academic subject matters [12]. Clickbait generally lures readers in by inciting their interest with a small glimpse of what the story is about, which is known as "forward referencing narrative strategy" [12]. In this manner, readers are persuaded to read the rest of the article and this increased traffic on a website enables the owners to raise the price of advertising [12].

Clickbait headlines tend to hold back some aspect of the story, leaving the reader feeling that they are missing a piece of information, therefore they are a very effective method of driving readers to specific websites [13]. Clickbait headlines usually send readers to a different website to the one they are on, or else keep them on the same website if it has internal articles advertised [13]. The popularity of a website is measured in terms of hits and visits, and this directly relates to the price of advertising on that website, so the main aim is to encourage traffic on a website [13]. Increasingly, readers can share the articles through their social media, displaying the clickbait headline for more people to see [13].

Clickbait headlines can be incredibly damaging, as often they do not encourage readers to view the entire article, and if they are interpreted on their own merit, they can perpetuate misunderstanding and warped truth [12]. For example, Jenkins wrote on the case of a young girl with facial disfigurement being asked to leave a fast-food chain due to her disfigurement [14]. This story was widely shared on social media with a clickbait headline. Although it transpired that this story was untrue, it was too late to reverse the negative impact on the brand. Another example of the power of clickbait is the 2008 account of Steve Jobs' heart attack, which caused an immediate drop, of around 10%, in Apple's stock price [15].

Clickbait headlines are not a widely researched topic, although the effect of news headlines in general on readers has been studied by various theorists, such as Kim et al. and Gibbons et al. [16] [17]. It has not been verified exactly how readers react to clickbait headlines and their sensationalised approach. Doyle stated that some internet users might view clickbait headlines as junk, potentially something harmful to their computer [18], which Tandoc suggests could risk reducing the worth of the online media if genuine articles are discounted [19]. Furthermore, Loewenstein affirms that if clickbait headlines are too over-the-top then users will simply not pay attention to them [9].

### 2.2. Forward referencing

A method often used by journalists to capture readers' interest is "forward referencing", which is, according to Blom and Hansen, the act of alluding to aspects that will occur further on in a story at an earlier point, often leaving things unexplained, vague or avoiding directly referring to a key feature of the story [20]. A study undertaken in Denmark by Blom and Hansen investigated 2000 headlines chosen from a news website at random, and from this they identified two commonly-occurring forms of forward referencing: deixis and cataphora [20]. Deixis refers to forward referencing at discourse level, for example "This news will blow your mind" whereas cataphora refers to forward referencing at phrase level, as with "This name

is hilarious" [13]. Unfortunately Blom and Hansen recommend no method for the automatic detection of such phrases, but they discovered they mostly crop up on websites that are commercial and reliant on the funding of advertising, or tabloid news [20].

Ifantidou carried out a study in which his participants were not bothered by vagueness in headlines; they were only concerned to have original, gripping headlines [21]. This study focused on college students, so it can be taken that in the case of such a demographic, forward referencing did not irritate, but incited their interest and intrigued them, as long as it was still deemed original and gripping [21]. However there have not been any studies to carry out a comprehensive analysis on the effect of forward referencing, and whether it is more likely to inspire intrigue or irritation [20].

### 2.3. The curiosity gap

Clickbait relies on the use of intriguing messages that open a "curiosity gap" in readers, requiring them to click on the article to fill that gap [13]. However, while clickbait headlines push readers to click on them, they often end up dissatisfied when the article does not intrigue them as much as the headline did [13]. The information-gap theory is widely employed in research on this phenomenon, offering a psychological

motivation [9]. A person becomes intrigued by enigmas in the story, unknown conclusions of an event, previously-known information that they have failed to recall, knowledge held by others but not them, or a diversion from what they anticipated to happen [22]. The information-gap theory explains how a reader will suddenly become aware of a gap in their expertise and be driven through inquisitiveness and a feeling of inadequacy to remedy that sentiment [23]. Studies have observed that clickbait headlines are purely a vehicle for procrastination, encouraging readers to go from article to article, eventually feeling weighed down with needless information and therefore unlikely to seek less superficial news stories. A further worry is that it is hard to monitor clickbait journalism in today's online media environment [22].

### 2.4. Clickbait, Use of Social Media and Data-Driven Optimisation

Clickbait is particularly successful due to the widespread use of social media. Websites such as Facebook and Twitter are ideal vessels for spreading clickbait stories, and can manipulate readers' experiences by using data to optimise the stories shown [13]. Furthermore, social media can provide immediate feedback for news providers. In the print media, receiving feedback for articles is a long, drawn-out process with ample opportunity for intervening factors to reduce the reliability of results that

could improve sales, however clickbait provides a quick method for measuring response [24]. The successful results achieved by clickbait on social media inspired Facebook to declare an attempt to cut down on clickbait in 2014, stating that it would endeavour to get rid of clickbait articles in its feeds [25]. Exactly what Facebook does in attempt to filter clickbait has not been made widely known, but it potentially incorporates aspects such as time spent on linked pages and the correlation between clicks and 'likes' [25]. Such factors are not very informative in the detection of clickbait as they heavily rely on context, thus many users of Facebook still note the prevalence of clickbait articles and Facebook has recently stated again that it will strive to reduce this [25]. A recent study observed how clickbait was used on Twitter [13]. However, as clickbait is a common feature not only of social media but across the internet, studies from such a specific angle are not necessarily the most effective method to investigate clickbait, although evidently social media has become so universal that the effect of clickbait on these websites is significant [26]. Clickbait is used on respectable websites such as The Washington Post and The Guardian also, which can increase their profits, but risks devaluating their websites, and reducing readers' views on their reputability [26]. Consequently, online media outlets might be advised to refine their approach to drawing in readers.

## 2.5. Clickbait and Eye Movement Tracking

Eye movement tracking is a practice that has been employed increasingly over the past ten years. It can provide a valuable insight into a person's visual and mental happenings [27]. Over the years, it has been used in fields such as analogy, reading, driving, image scanning and arithmetic, amongst others [27]. Development of this practice has identified various important factors that are considered to be strong signs of visual behaviour, such as fixations, pupil dilation, scan paths and saccades [28]. Eye fixations tend to signify that information is being obtained and understood, so they are particularly relevant to this study; this is when the eyes hold a steady stare at a precise spot on the observation area, lasting from 200-300 milliseconds [28]. Pupil dilation, on the other hand, tends to represent a person's level of attentiveness and titillation in the object they are looking at [28]. Researchers most commonly use fixations and saccades (which are defined as sharp movements between fixations) in eye-movement tracking; often they observe optical activity such as the length of fixations, the transitions of fixations, or the speed and size of saccades [28].

Studies such as Bradley et al., Stanners et al. and Partala et al. have confirmed that the eye's pupils dilate when a person is witnessing something they find either particularly disagreeable or particularly pleasing

[29] [30] [31]. Pengnate studied internet users' level of titillation and consequent actions when reading news and clickbait headlines online by employing eye movement tracking [12]. Pengnate's research used an unprecedented method, examining the emotive response of a person through subtle changes in their ocular movements, particularly in regards to pupil dilation [12]. In Pengnate's study, respondents' pupils were expected to expand when they were titillated by the headlines that they read [12]. His findings were a valuable influence on the study of clickbait and its effect, as he discovered a correlation between titillation and the drive to read more of a news story. Desmet has affirmed that in such studies, it is more valuable for research to document bodily indicators of response than to ask respondents for their own views [32]. Not only does this take away the consideration of a respondent's emotional response, but it creates less interruption during the process of study and allows for variations in respondents' reactions to be registered immediately [32]. The method, however, was imperfect and faced restrictions. Firstly, eye movement tracking technology is expensive, cumbersome and laborious, as explained by Masiocchi and Still, however in future study these limitations might be reduced, as cheaper and more efficient equipment is in development (for example by grinbath.com) [33]. A second limitation, as expressed by Desmet, is that examining pupil dilations only gives an overview of a

response, registering a level of stimulation but not specific thoughts, reactions or feelings [32].

## 2.6. Clickbait and Business

### *2.6.1. How does clickbait generate profit?*

Although paid subscriptions are being introduced by more and more by websites, advertising remains the biggest contributor to overall financial gain [34]. Internet media websites rely increasingly on advertisement for profit [35]. Hits and clicks are incredibly important for the owners of such websites to draw in advertisers and to negotiate better deals, therefore clickbait is aimed to make money by increasing website traffic [36]. As mentioned in the Introduction, some websites even pay writers extra for a high volume of clicks, as does Slant magazine.

Advertising online is known as a ‘pay-per-click’ or a ‘click-through-payment’ programme, whereby the website will be paid once an advertisement has been clicked on; the website owner agrees to advertise a product and receive payment only when a user is taken to the website of the advertiser [37]. This model highly encourages the use of clickbait as a strategy. Whereas some websites need advertising to continue financing themselves, others make the choice to be a place that visibly endorses goods, but in both of these scenarios plenty of visitors on the website are

required in order to generate revenue through pay-per-click [37]. Often the pay-per-click programme is moderated by a third-party organisation [38].

The pay-per-click programme tends to be associated with Internet giants such as Google. Most search engines, when a user has entered words into the search-box, generate the results showing sponsored links at the top [39]. The sponsored advertisements that appear in a Google search are subscribing to the pay-per-click model, but with modifications: most likely an advertiser has paid extra for the prime position as the first link that appears with search results [39]. Advertisers compete against one another for this spot, and must make additional payment on top of following the standard pattern for pay-per-click payment. Google is comparably open about paid advertisements; some websites do not inform users of this, trying to conceal the fact that they are showing links that have paid extra to be more visible. Of course, with a pay-per-click advertising model the website needs a large volume of users to click on the advertisement, and this leads to the advertisements themselves presenting clickbait headlines [40]. Some clickbait links direct users to advertisement webpages which request that they sign up; Guinness noted an experience such as this, where selecting the choice not to sign up only redirected him to another website, with additional advertisements [40].

Overall views and popularity of websites that display clickbait headlines are also key elements, even without considering pay-per-click advertising. Websites with more hits are able to ask for higher fees for advertisement. Generally, clickbait headlines do not advertise products; the objective is to increase revenue gained from advertising by luring readers onto the website, and encouraging them to share such stories on their social media, to ultimately gain more website traffic [41].

### *2.6.2. The Buzzfeed approach*

Buzzfeed is a good case study of a company that enjoys success but has been criticised for use of clickbait, and for repurposing of information from other articles. Buzzfeed, established in 2006 with rapid growth since then, hosts contents on many subjects, including entertainment and general news, with articles, videos, quizzes and many other varieties of content [42]. In 2013, Buzzfeed was estimated to have peaked as a pioneering new business, and established itself globally in that year also [42]. Buzzfeed is now worth $1 billion, or at least that was the price paid for ownership by Disney in 2014, despite the fact that Buzzfeed was only worth $1 million in 2012 [42]. In February of this year, Buzzfeed's worth was estimated to be around $1.6 billion, encouraging spectators to ask what this website is doing right, enjoying much greater success than well-known established news websites [42].

Buzzfeed was designed with social media sharing in mind and its content is constructed with the aim of facilitating sharing on social networks [43]. Whilst the general pattern for such websites is to show advertisements through pop-ups or in banners, Buzzfeed's founder had other aspirations. He wanted his website to be a vibrant viral outlet with constantly increasing webpage hits, not only a source of entertainment and news but a successful model for simplified content sharing [43]. He formulated a pattern, beginning with the use of catchy headlines to entice a reader to click on a link, which then refers the reader to a page of content, showing advertisements alongside or within. The next stage of this pattern would be for the reader to share the page through their social networks, spreading the content to their online connections. This then results in increased views of the website, and the advertisements will be shared and viewed as often as the article through the likes of Facebook, Twitter and, more recently, Snapchat [43]. Buzzfeed's provision of shareworthy content and the widespread use of social media websites enabled it to build its huge readership, according to the Business Insider's CEO/Editor Henry Blodget [43]. Blodget, however, advocates an approach that is not so heavily reliant on social media, as he considers it to be rather over-emphasised in modern online practice [44]. This is because social media is just one small way that websites can raise their advertising profit; Business Insider makes use of multiple tactics in trying to increase website hits and advertisement links,

by ensuring its content is highly visible on search engines, email and social media alike [44]. Blodget recommends that all three of the above methods should be exploited to entice a large, and varied, readership [44].

Furthermore, Buzzfeed takes an innovative approach to advertising, mixing its content with advertisement in a specially-designed format [45], rather than keeping these two things separate. This method been proven to surpass traditional online method of advertising [45]. Dan Frommer has explained how Buzzfeed maximises its use of advertisement. One method is to display advertisements in a format that showcases a brand or product within content, a poll in Frommer's example. Another method is to showcase an advertisement as if it were a piece of content on its own, yet once a reader has clicked on the link, advertisement videos appear instead of an article as expected [45]. The success of these models could be attributed to the fact that advertising in what is called a 'story unit' is more effective in getting clicks than a search engine advertisement. Traditional advertisement methods receive a click rate which is on average from 0.25% to 0.7% [45], whereas Figure 2 displays that whilst Google has a click rate of roughly 1%, advertisements such as Buzzfeed's in the 'story unit' approach have a 1.5% to 5% click rate [45].

**Figure 2. A click through rate comparison between two types of advertisement [45]**

# CHAPTER 3

# Methodology

Clickbait is defined as '(On the Internet) content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page [11].' It also refers to the links to such sites: the 'bait' which attracts visitors to the site. That clickbait works seems self-evident: otherwise it would not exist. In this study, we aimed to determine how 'clickbait' works and how people react to different kinds of clickbait. The objective of this project was to investigate that. Buzzfeed.com was chosen to test the participants' browsing performance because it is famous for providing clickbait headlines. This was mainly a qualitative approach. We wanted to observe if people react to clickbait – and which kinds they are attracted to. We also wanted to observe demographic trends (for example, males are attracted to different clickbait than females), however, we have no a priori hypotheses about this.

## 3.1. Design

As this study aimed to discover people's reactions to different kinds of clickbait, the experiment was designed to record the participants' real browsing performance when they are shown clickbait. Participants were invited to come to a computer lab and they were required to use an eye

tracker to do this experiment. The reasons for selecting this method are discussed in the following subsections. First, the participants were provided with an information sheet (Appendix B) which gave them information about project title, what will happen in this experiment, time commitment, participants' rights, benefits and risks, reimbursement, confidentiality and contact information. The participants were then required to sign a consent form (Appendix A) to confirm that they agreed to join this experiment.

The participants were asked to imagine that they had 15 minutes to spare and to spend that time surfing the web for amusement. They were given a page from which to start, which happens to contain a lot of clickbait (https://www.buzzfeed.com/). They were told that their interactions with the browser would be recorded and their eye-movements tracked, however, they were not told that the study was specifically focusing on clickbait because it might influence their clicking decisions and browsing performance. They were left alone to do this experiment. After they thought they had completed 15 minutes (we were interested in their perception of time) or after 20 minutes (whichever is sooner) they were stopped. Then, in order to collect demographic information, the participants were provided with a short questionnaire (Appendix C) which contained questions regarding their age, gender, occupation, education level and browsing habits. After they finished the questionnaire, they were shown a replay of

their interaction (mouse movements and clicks and eyetracks) and asked to recall what they were thinking. For instance, 'Why did you click on that?'; 'Why did you look at that but not click on it?' and ‘Were you attracted by the headlines or pictures?’ This interview was audio recorded. Those interactions and demographic information were analysed jointly to discover the relationship between demographic trends and different reactions to different kinds of clickbait. Finally, the author gave every participant a debriefing (Appendix C). It was then explained to them that we were interested in their reaction to the clickbait.

### 3.2. Participant

There were 21 participants involved in this experiment, including one participant for the pilot study. There were 9 women and 12 men. The participants’ ages range from 21 to 60 with a median age of 24. In detail, 14 participants were between 21 to 25 years old and three participants were between 26 to 30 years old and four participants were over 35 years old. The participants have diversified language backgrounds: six of them use English as their first language and 12 participants’ mother language is Chinese. There were three participants who speak French, Arabic and Spanish as the mother tongue. Because of the constraints on this experiment such as University’s regulations on borrowing the eye tracker from the University of York and using the computer lab, all of the

participants are staff or students from the University of York. The participants were recruited by sending an e-mail to their personal contacts. As an incentive, the participants were offered the chance to be entered in a random draw for one of two Amazon vouchers for £20.

### 3.3. Pilot study

A pilot study must be conducted to determine the feasibility of an approach before it is applied to the formal experiment. This pilot study was supervised and participated in by Dr. Alistair Edwards. All the processes and details of the pilot study were the same as the original design of the experiment.

After conducting this pilot study, Dr. Alistair Edwards and the author identified four modifications that needed to be made. First, the strategy for recording the participant's browsing activities was changed from 'Web recording' to 'Screen recording.' The eye tracker software system provides several different strategies to record the user's clicking activities and eye movements. 'Web recording' was originally selected because the experiment was based on a website. However, during the replay session, each click to a new page generated a new recording file. This made the replay session inefficient and it was difficult to observe the headlines that a participant clicked. Using 'Screen recording' solved that problem.

Second, we used a fixed seat instead of a roller seat because we found

that the machine would lose track of eye movement if the participant moved around during experiment. Moreover, an additional sentence was added to the information sheet requesting that the participant not move around during the experiment.

Third, the layout of the questions in the debriefing sheet was modified and improved. Finally, we selected Buzzfeed.com as the website being used in the experiment instead of Lifehacklane.com. Although the pictures and headlines in Lifehacklane.com contained some types of clickbait, most of them were about health and life tips. This experiment aimed to find people's reactions to different types of clickbait and Buzzfeed.com was selected due to its comprehensiveness.

The pilot study provided modifications to improve the feasibility of the experiment and guide the formal experiment. However, due to the modifications to the original design, the data from the pilot study was not used in the final analysis.

### 3.4. Buzzfeed

As an Internet media company, Buzzfeed defines itself as a 'social news and entertainment company' and claims that it provides original reporting, breaking news, videos, entertainment information and other interesting fields of knowledge [46]. Unlike the original choice of website, Lifehacklane.com, Buzzfeed covers a variety of topics such as news

(breaking news, celebrity news, and political news), quizzes, DIY and health. Buzzfeed has gained a large amount of page views in those areas. In 2015, Buzzfeed video's YouTube account obtained 9.3 million subscribers and 6.3 billion page views [47] and a post by a Buzzfeed editor regarding a debate over clothing colour garnered more than 28 million page views in one day [48].

According to Newswip, Buzzfeed was the top ranked of 'Facebook Publisher Rankings' from 2013 to 2014 [49]. However, according to the Pew research centre, Buzzfeed was treated as an unreliable source by most people in United States in 2014 [50]. One of the reasons is that many of Buzzfeed's headlines are criticised for their reference to clickbait. For example, according to the Advertising Standards Authority, an article with the headline "14 Laundry Fails We've All Experienced" was thought to be breaking UK advertising rules because its content promoted the brand 'Dylon,' however, the headline failed to 'make it clear that the main content of the article was about Dylon's online advertising [51]. This was supported by our pilot study which discovered that Buzzfeed still contained a lot of clickbait under different categories (news, videos and health tips). Finally, the supervisor Dr. Alistair Edward and the author decided to use Buzzfeed.com in the formal experiment.

### 3.5. Hawthorne effect

The Hawthorne effect describes how participants change and improve

their behaviours when they know that they are observed in an experiment [52]. Originally, this term was created by Henry A. Landsberger who found that worker productivity was improved if the workers worked under higher levels of light [53]. The point is that the very fact that a participant is being observed may affect their performance. This experiment aimed to observe the participants' browsing performance and their natural reactions to different kinds of clickbait. However, as it is an experiment, the author cannot recreate the participant's private browsing situation and the Hawthorne effect was present. For example, a clickbait which contains sexy pictures may attract people to click when they are in their private space but they will not click such a link when they are aware of being observed. This effect cannot be entirely avoided in a human-based experiment, however, its negative influence can be decreased.

According to Mcbride, there are two ways to deal with the Hawthorne effect: the first one is to use a naturalistic observation technique, which means to observe people unobtrusively; the other is to inform the participant that their identification information will be anonymous and make it clear that all of the data will be confidential and stored securely [54]. Considering these two techniques, the author chose to leave the participant alone when they started browsing the website in order to create a naturalistic situation and clearly informed them about the confidentiality policy in the information sheet. Furthermore, this study

used an eye tracker to record the participants' performance and the author would ask questions if a participant spent a lot of time looking at a link but did not choose to click. This technique might also eliminate some performance bias which was caused by the Hawthorne effect.

Note, however, that the participants were briefed and were always aware that their actions and eye movements were being recorded. This was necessary ethically, however, it may have meant that they did not behave naturally.

### 3.6. Eye movement

In order to get a full picture of the participants' behaviour, as well as recording their interaction with the browser, the participants' eye movements were tracked. In particular, this was intended to capture their natural reactions to see whether their browsing behaviour was affected by the knowledge that they were being observed. This might be a particular instance of the Hawthorne effect, whereby clickbait has succeeded in attracting the participant's attention (as shown by their eye gaze turning to the clickbait) and yet it fails inasmuch as the person chooses not to click on it. It might be that the person is interested in the clickbait but in an experiment they would be embarrassed about admitting so. Specifically, some clickbait is sexually provocative and some participants might prefer to not admit that it has attracted their attention.

The participants were required to use a SMI laptop to browse a website and their eye movements were recorded by an eye tracker called "RED 250 Mobile." In order to increase accuracy of the eye tracker, the participants were asked to find their most comfortable positions and conduct the calibration test. Their chairs were designed to be fixed and hard to move. After the eye movement recording, the author used the SMI BEGAZE software to replay the recordings and analyse results.

### 3.7. Ethics

A minor deception was that the participants were not told in advance that we were specifically looking at their reaction to clickbait. This was because their consciousness would influence their performance and clicking decisions. The information about clickbait was revealed in the debriefing on completion of the session. (At which point and before the replay they may choose to withdraw).

Another potential problem was that some clickbait used sexually provocative pictures to attract attention. The participants may be embarrassed if it is shown that their attention was thus captured. However, knowing that they are being recorded, they may not click on clickbait that they would have done in privacy.

The former problem can only be dealt with sensitivity. That is, the participants will not be forced to answer why they did or did not react to

some items. The latter problem is more related to methodology; it was discussed in previous sub sections that this lab-based experiment may not fully reflect real-world behaviour because of the Hawthorne effect.

Finally, as suggested by Dr. Alistair Edwards, these ethical concerns were clearly stated in the ethical form and this form was approved by the Physical Science Ethics Committee. According to the feedback from Ethics Committee, we made a clear statement in the information sheet (Appendix B) that sexually provocative materials might be encountered. In addition, we provided an informed consent form (Appendix A) for each participant before the experiment in order to make it clear that they read and understand the potential ethical issues.

### 3.8. Generating Clickbait Categories

After conducting this experiment, the records showed that 151 links of headlines with images were clicked by 20 participants. Our analysis was based on those headlines and images using the qualitative categorisation method. There were three main sections: 1. Code all 151 links and analyse it by considering different groups such as gender and age; 2. Create another coding scheme to separate the 151 links into: Non-Clickbait, Normal Clickbait and Pure Clickbait; 3. Analyse the data based on Normal Clickbait and Pure Clickbait; compare the results with section 1.

For the first coding, although Buzzfeed has some categories for

different headlines, those categories are not specific enough. Thus, the author chose to use emergent codes and create a set of categories from those click results. We took a subset of those headlines and generated a set of categories which included all of the valuable elements in the headlines. Then the author sought the supervisor's help to modify the codes and generate a consolidated list. Finally, the coding scheme was completed and shown as below. According to the data, 19 labels were created such as 'A list,' 'Health' and 'Political news.'

| No. | Label | Description |
|---|---|---|
| 1 | A list | Headlines or images include the element of a list, such as '10 amazing photos you've never seen.' |
| 2 | Game | Headlines or images include the element of game. |
| 3 | Health | Headlines or images include the element of health tips or lifestyle. |
| 4 | Food | Headlines or images include the element of food. |
| 5 | Travel | Headlines or images include the element of travel. |
| 6 | Celebrity | Headlines or images include the element of celebrity. |
| 7 | Political news | Headlines or images include the element of political issues. |
| 8 | Accident news | Headlines or images include the element of an accident. |
| 9 | A new or unknown concept | Headlines or images include the element of a new or unknown concept. |
| 10 | Animals | Headlines or images include the element of cute animals. |
| 11 | Makeup | Headlines or images include the element of makeup information. |
| 12 | Shopping | Headlines or images include the element of shopping. |
| 13 | Dress | Headlines or images include the element of dress. |

| | | |
|---|---|---|
| 14 | Social media news | Headlines or images include the element of social media news. |
| 15 | You or I | Headlines or images include the element of 'You' or 'I.' |
| 16 | Quizzes | Headlines or images include the element of quizzes. |
| 17 | Life sharing story | Headlines or images include the element of sharing people's own life. |
| 18 | Education | Headlines or images include the element of education |
| 19 | Advertisement | Headlines or images include the element of advertisement |

**Table 1. First coding scheme of 151 headlines with images**

For the second coding, the main purpose was to detect how many of the headlines were clickbait. The process of this categorisation was the same as the first coding. The labels and descriptions were listed as below.

| No. | Label | Description |
|---|---|---|
| 1 | Non-clickbait headlines | The headline objectively stated the information without a sensational or provocative nature. |
| 2 | Clickbait headlines | 'The headline whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page' [11] |
| 3 | Pure clickbait | The headline itself is an advertisement and links to an advertising website |

**Table 2. Second coding scheme of 151 headlines with images**

### 3.9. Inter-coder reliability

In order to establish that the coding is reliable, it is necessary to perform an inter-coder reliability step. Theoretically, there are two different methods to select the second coder. The first method is to select an insider coder, which means that the second coder is an insider of this research

method and very familiar with the terminologies used in this study. However, the insider coder would be constrained by this familiarity and the results of inter-coder reliability would not be precise. For example, two coders who have very similar knowledge backgrounds and thinking methods may not detect problems within coding scheme because 'standers-by see more than gamesters.' The second method is to recruit an outsider coder. An outside coder make not fully comprehend the information of his field, however, they may provide more objective opinions. In this study, the author invited an outside coder to optimise the objectivity of coding. 20 headlines were randomly selected to test inter-coder reliability. Because two coders might agree by chance, the author chose to use a sophisticated measure of inter-coder reliability which is Cohen's Kappa. The formula is:

$$\mathbf{K = (Pa-Pc)/ (1-Pc)}$$

Where Pa equals the percentage of cases where the coders agree and Pc equals the percentages of cases where the coders would agree by chance

For the first coding, there are three disagreements between the two coders, therefore, Pa equals 0.85 (Pa= (20-3)/20=0.85). As there are 19 labels, Pc equals 0.053 (Pc= (1/19*1/19)*19=0.053). Therefore, the Cohen's Kappa is (Pa-Pc)/ (1-Pc) = (0.85-0.053)/ (1-0.053) = 0.84. Because the Cohen's Kappa is over 0.80, the agreement between the two

coders is excellent and the initial coding scheme is comprehensible and clear.

For the second coding, there are two disagreements between the two coders, therefore, Pa equals 0.90 (Pa= (20-2)/20=0.90). As there are three labels, Pc equals 0.053 (Pc= (1/3*1/3)*3=0.33). Therefore, the Cohen's Kappa is (Pa-Pc)/ (1-Pc) = (0.90-0.33)/ (1-0.33) = 0.84. Because the Cohen's Kappa is over 0.80, the agreement between the two coders is excellent and the coding scheme for the first coding is comprehensible and clear.

# CHAPTER 4

# Result

## 4.1. Section 1: Analysis based on coding of all the 151 headlines with images (First coding)

In this section, the results are based on all 151 headlines with images and consider different gender and age groups. The occurrences and percentages of 19 coding labels were shown in Table 3 and Figure 3. First, 'A list', 'Political news' and 'Game' were the three most popular categories with 16.3%, 15.8% and 15.3% respectively. 10.5% of headlines have an element of 'Celebrity'. Following that, 9.1% of headlines contained a 'You and I' element, 8.6% contained information about 'Quizzes', 4.3% contained references to 'Health' and 3.8% of headlines were about 'A new or unknown concept.' Moreover, both 'Food' and 'Travel' accounted for 2.9% of the headlines and 'Animals,' 'Makeup' and 'Life sharing story' accounted for 1.9% each. Furthermore, 'Shopping' and 'Dress' both occurred three times and accounted for 1.4% of the headlines. Only 0.5% of the headlines mentioned 'Accident news,' 'Social media news,' 'Education' and 'Advertisement.'

| Coding | Occurrence | Percentage |
|---|---|---|
| Accident news | 1 | 0.5% |
| Social media news | 1 | 0.5% |
| Education | 1 | 0.5% |

| | | |
|---|---|---|
| Advertisement | 1 | 0.5% |
| Shopping | 3 | 1.4% |
| Dress | 3 | 1.4% |
| Animals | 4 | 1.9% |
| Make up | 4 | 1.9% |
| Life sharing story | 4 | 1.9% |
| Food | 6 | 2.9% |
| Travel | 6 | 2.9% |
| A new or unknown concept | 8 | 3.8% |
| Health | 9 | 4.3% |
| Quizzes | 18 | 8.6% |
| You or I | 19 | 9.1% |
| Celebrity | 22 | 10.5% |
| Game | 32 | 15.3% |
| Political news | 33 | 15.8% |
| A list | 34 | 16.3% |

**Table 3. Overview of coding data with occurrence and percentage**

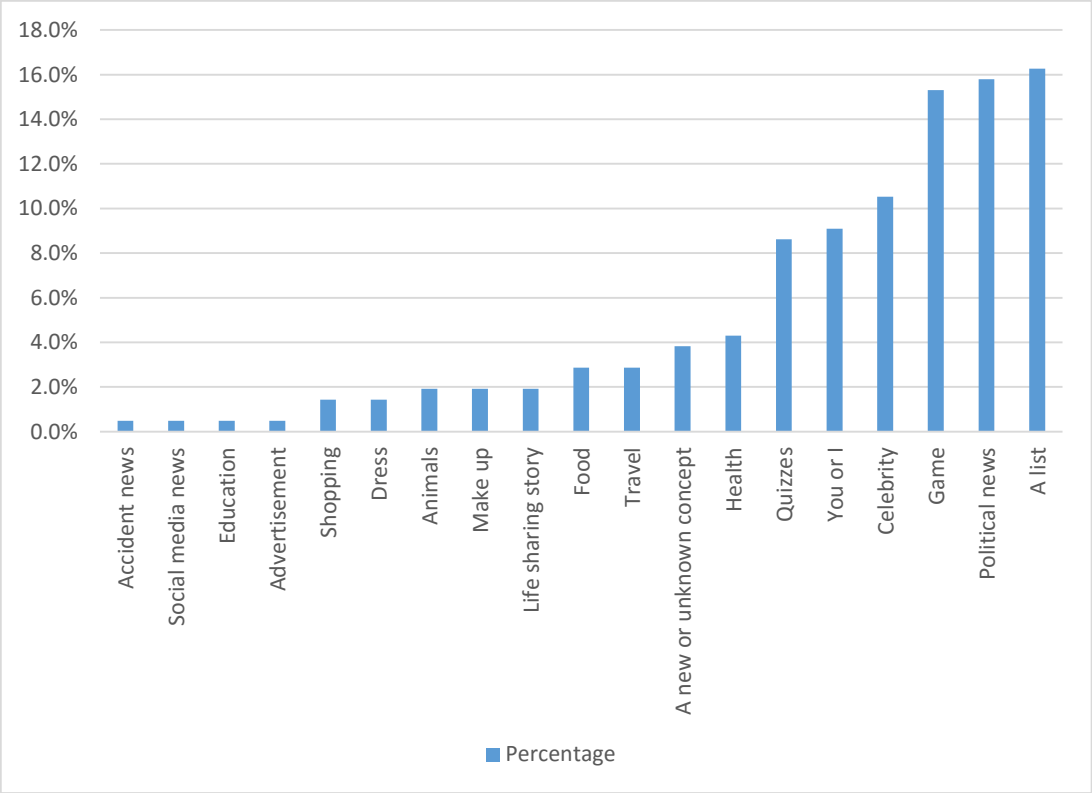


**Figure 3. Percentage of Headlines that contain these 19 Coding Elements**

Based on the above results, the author compared the click preferences of males and females. The gender with the higher percentage of click

preferences were marked in bold in table two. Females had a decentralised preference distribution, covering all 19 coding elements. In contrast, the headlines which were clicked by males covered only 10 elements, which were only half of the coding elements. Specifically, the headlines which were clicked by males did not contain any elements of ‘Food,’ ‘Accident news,’ ‘Animals,’ ‘Makeup,’ ‘Dress,’ ‘Social media news,’ ‘Education’ or ‘Advertisement.’ However, 15.12% of headlines clicked by female contained those elements and their percentages ranged from 0.72% to 4.32%. The top three coding elements for males were ‘Game,’ ‘A list’ and ‘Political news’ with 20%, 17.14% and 15.71% respectively. The top three coding elements for females were ‘Political news,’ ‘A list’ and ‘Game’ with 15.83%, 15.83% and 12.95% respectively. The two gender groups had very similar results for the coding element ‘shopping,’ which were 1.43% and 1.44%. For the remaining elements, females had the higher percentages in the coding elements ‘life sharing story,’ ‘health,’ ‘quizzes’ and ‘political news.’ Males had the higher percentages in coding elements of ‘A new or unknown concept,’ ‘Travel,’ ‘You and I,’ ‘Celebrity,’ ‘A list’ and ‘Game.’ However, most of these differences were small except for ‘Food’ (4.32% difference), ‘Travel’ (6.42% difference) and ‘Game’ (7.05% difference).

| Coding | Male | Female |
|---|---|---|
| | Percentages | |
| Food | 0.00% | 4.32% |
| Accident news | 0.00% | 0.72% |
| Animals | 0.00% | 2.88% |
| Make up | 0.00% | 2.88% |
| Dress | 0.00% | 2.16% |
| Social media news | 0.00% | 0.72% |
| Education | 0.00% | 0.72% |
| Advertisement | 0.00% | 0.72% |
| Shopping | 1.43% | 1.44% |
| Life sharing story | 1.43% | 2.16% |
| Health | 2.86% | 5.04% |
| A new or unknown concept | 5.71% | 2.88% |
| Travel | 7.14% | 0.72% |
| Quizzes | 7.14% | 9.35% |
| You or I | 10.00% | 8.63% |
| Celebrity | 11.43% | 10.07% |
| Political news | 15.71% | 15.83% |
| A list | 17.14% | 15.83% |
| Game | 20.00% | 12.95% |

**Table 4. Percentage of different click preferences between male and female**

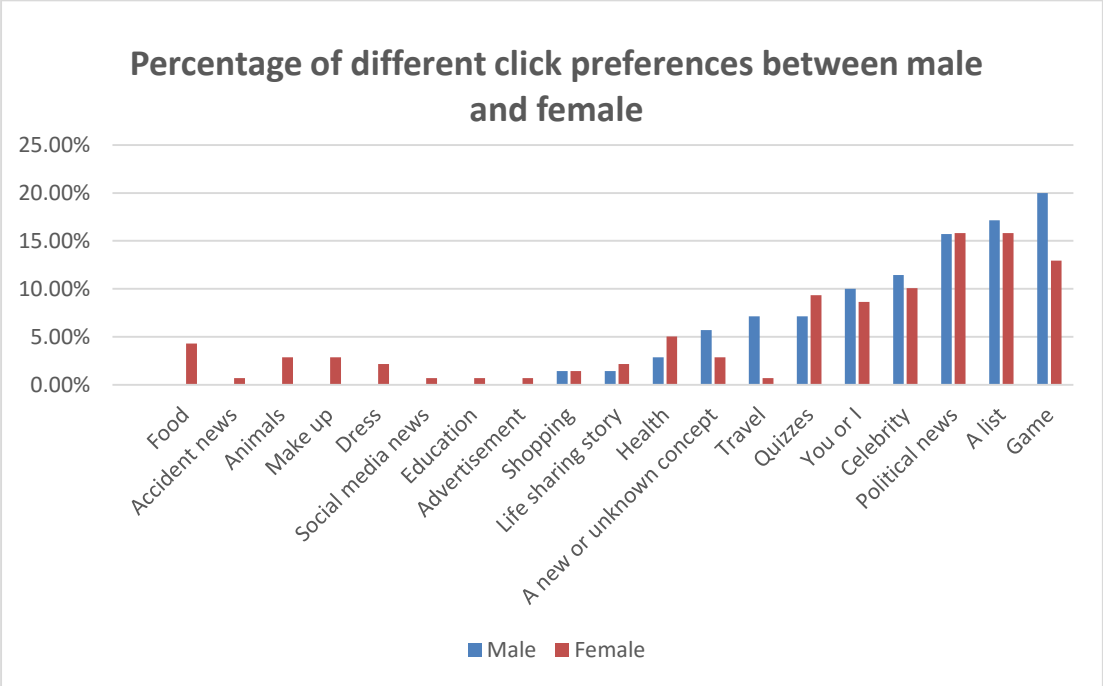


**Figure 4. Percentage of different click preferences between male and female**

The author compared the click preferences between people under 30 years old and people over 30 years old, which is shown in Table 5 and Figure 5. The age group with the higher percentage click preferences is marked in bold in Table 5. The young group had a decentralised preference distribution, covering all 19 coding elements. In contrast, the headlines which were clicked by old group covered eight elements, which is less than half of the coding elements. The top three coding elements for the young group were 'Game,' 'A list,' 'Celebrity' and 'Political news' with 17.68%, 16.57%, 10.50% and 10.50% respectively. The top three coding elements for the old group were 'Political news,' 'A list,' 'Celebrity' and 'Travel' with 50%, 14.29%, 10.71% and 10.71% respectively. Specifically, the old group did not click any headlines coded 'Game,' 'Food,' 'Accident news,' 'Animals,' 'Makeup,' 'Shopping,' 'Dress,' 'Social media news,' 'Quizzes,' 'Education' and 'Advertisement.' In contrast, 40.87% of headlines clicked by young people contained those elements and the percentages for them ranged from 0.55% to 17.68%. Furthermore, the young group had higher percentages in coding elements of 'Health,' 'A new or unknown concept,' 'You or I' and 'A list,' which were 4.42%, 3.87%, 9.94% and 16.57% respectively. For the old group, the results concentrated on 'Political news,' 'A list,' 'Celebrity' and 'Travel.' This group had higher percentages in the coding elements: 'Political news,' 'Celebrity,' 'Travel' and 'Life sharing story.' The most obvious differences for these two groups were elements

of ‘Political news’ (39.5% difference), ‘Game’ (17.68% difference), ‘Quizzes’ (9.94% difference) and ‘Travel’ (9.05% difference).

| Coding | Under 30 years old | Over 30 years old |
|---|---|---|
| | Percentages | |
| Game | 17.68% | 0.00% |
| Food | 3.31% | 0.00% |
| Accident news | 0.55% | 0.00% |
| Animals | 2.21% | 0.00% |
| Make up | 2.21% | 0.00% |
| Shopping | 1.66% | 0.00% |
| Dress | 1.66% | 0.00% |
| Social media news | 0.55% | 0.00% |
| Quizzes | 9.94% | 0.00% |
| Education | 0.55% | 0.00% |
| Advertisement | 0.55% | 0.00% |
| Health | 4.42% | 3.57% |
| A new or unknown concept | 3.87% | 3.57% |
| You or I | 9.94% | 3.57% |
| Life sharing story | 1.66% | 3.57% |
| Travel | 1.66% | 10.71% |
| Celebrity | 10.50% | 10.71% |
| A list | 16.57% | 14.29% |
| Political news | 10.50% | 50.00% |

**Table 5. Percentage of different click preferences between two age groups**

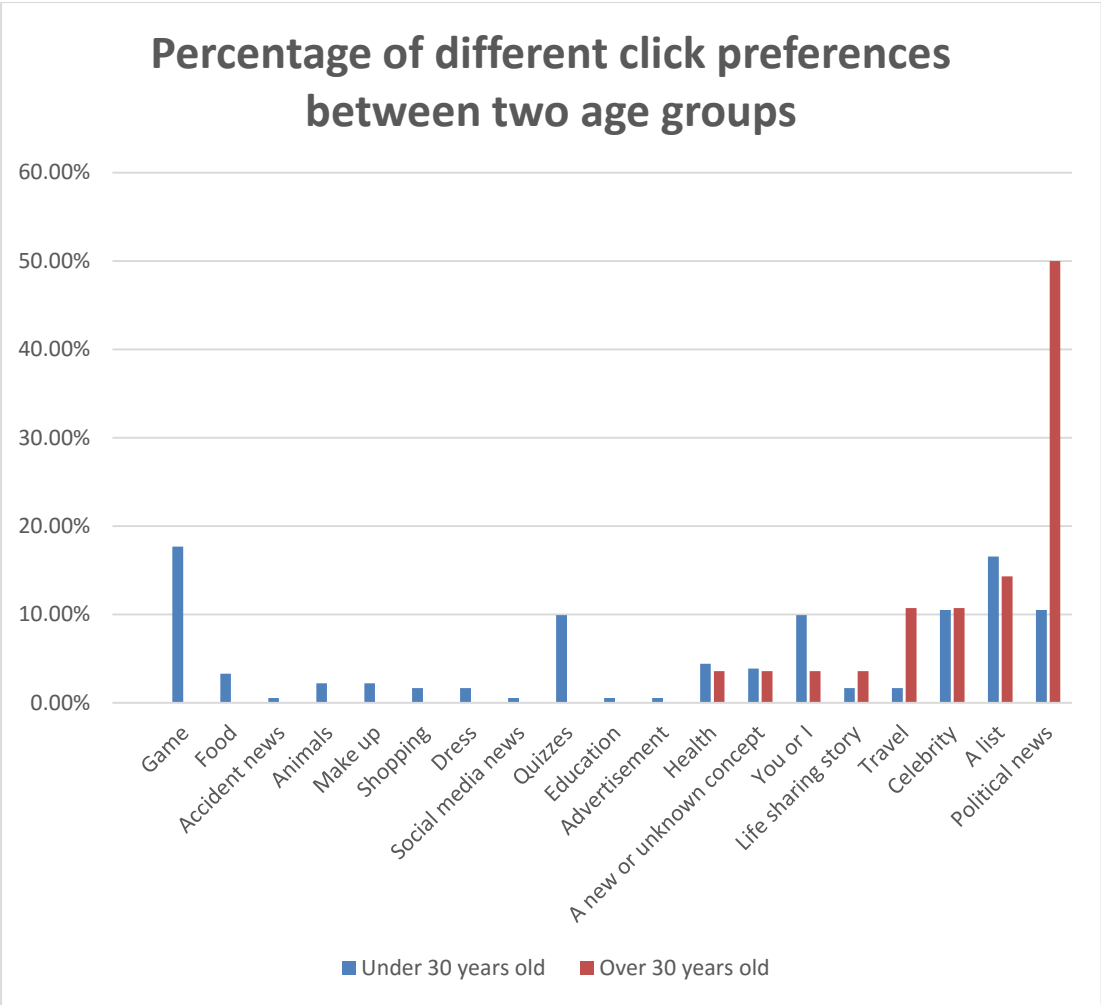


**Figure 5. Percentage of different click preferences between two age groups**

### 4.2. Section 2: Detect Clickbait (Second coding)

The main purpose of Section 2 is to detect clickbait headlines. The second coding scheme was used to divide the headlines into: non-clickbait headlines, normal clickbait headlines and pure clickbait headlines (direct advertising clickbait). According to Table 6 and Figure 6, 15.89% of headlines were categorised as non-clickbait headlines, 83.44% of headlines were defined as normal clickbait and only 0.66% of headlines were pure clickbait.

| Coding | Occurrence | Percentage |
|---|---|---|
| Non-Clickbait | 24 | 15.89% |
| Normal Clickbait | 126 | 83.44% |
| Pure Clickbait | 1 | 0.66% |

**Table 6. Overview of the second coding result**

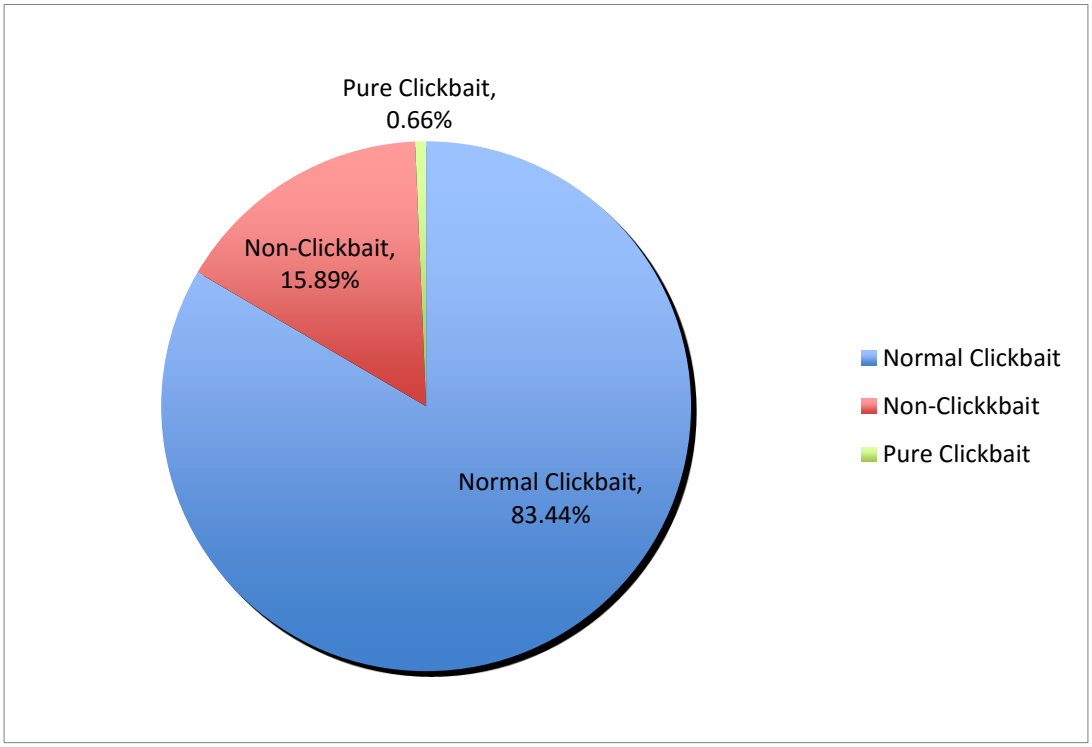


**Figure 6. Percentage results of Non-clickbait, Normal Clickbait and Pure Clickbait**

The author also calculated the number of coding elements (from coding scheme 1) per headline for non-clickbait and clickbait headlines. According to Table 7 and Figure 7, 70.80% of non-clickbait headlines contained only one coding elements (such as 'Political news'). None of this type of headlines contained three or four elements. There were two elements in the remaining 29.20% of headlines. In contrast, the percentage

of clickbait headlines that contained two elements increased to 42.5%. 15% of clickbait headlines included three coding elements and 0.8% of clickbait headlines had four coding elements.

| Number of coding elements per title | Non-clickbait headlines | Clickbait headlines |
|---|---|---|
| | Percentage | |
| 1 element | 70.80% | 41.7% |
| 2 elements | 29.20% | 42.5% |
| 3 elements | 0.00% | 15.0% |
| 4 elements | 0.00% | 0.8% |

**Table 7. Comparison of Non-clickbait and clickbait headlines based on number of coding elements per title**

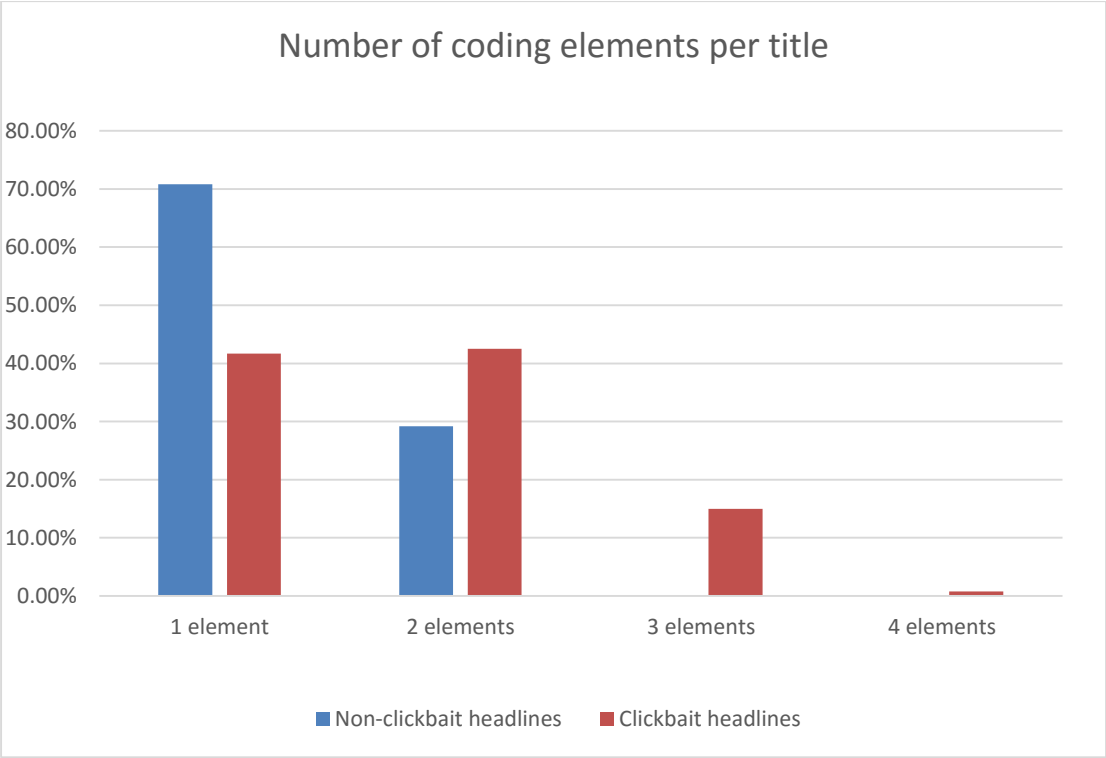


**Figure 7. Comparison of Non-clickbait and clickbait headlines based on number of coding elements per title**

### 4.3. Section 3: Analysis based on clickbait headlines considering different groups

This section generated results from clickbait headlines and the main difference between section 1 and section 3 is that section 3 removed all the non-clickbait data. According to Table 8 and Figure 8, 'A list,' 'Game' and 'Celebrity' were the top three elements of clickbait headlines. Compared with section 1, most percentages increased, however, the percentages for 'Accident news', 'Life sharing story' and 'Political news' decreased after removing non-clickbait data, particularly 'Political news,' which decreased from 15.80% to 7.10%.

| Coding labels | Section 1 | Section 3 |
|---|---|---|
| | Percentage | |
| Accident news | 0.50% | 0.00% |
| Social media news | 0.50% | 0.50% |
| Education | 0.50% | 0.50% |
| Advertisement | 0.50% | 0.50% |
| Shopping | 1.40% | 1.60% |
| Dress | 1.40% | 1.60% |
| Life sharing story | 1.90% | 1.60% |
| Animals | 1.90% | 2.20% |
| Make up | 1.90% | 2.20% |
| Food | 2.90% | 3.30% |
| Travel | 2.90% | 3.30% |
| A new or unknown concept | 3.80% | 3.80% |
| Health | 4.30% | 4.90% |
| Political news | 15.80% | 7.10% |
| Quizzes | 8.60% | 9.80% |
| You or I | 9.10% | 10.30% |
| Celebrity | 10.50% | 10.90% |
| Game | 15.30% | 17.40% |
| A list | 16.30% | 18.50% |

**Table 8. Overall comparison between Section 1 and Section 3**

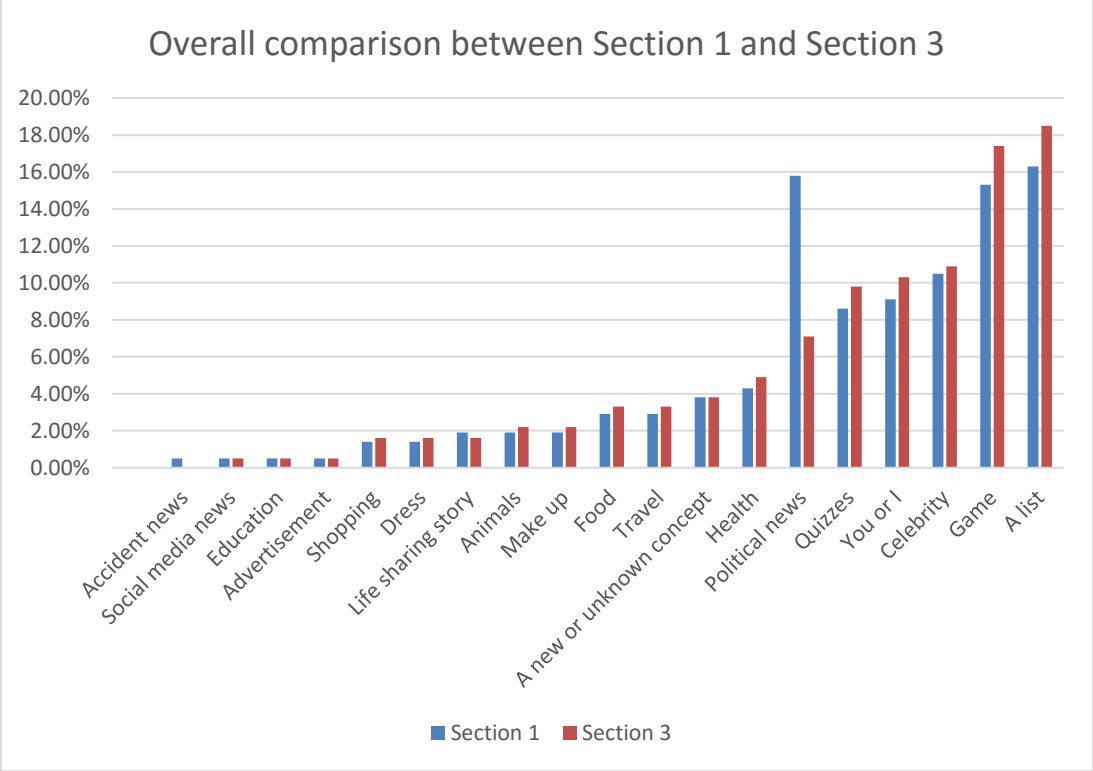


**Figure 8. Overall comparison between Section 1 and Section 3**

For different gender groups, the biggest difference was that percentage of 'Political news' decreased considerable for both genders. Consequently, 'Political news' was not one of the top three elements for either group. The male group preferred headlines and images that contained elements of 'Game' (23.73%), 'A list' (20.34%), 'You and I' (11.86%) and 'Celebrity' (10.17%). Females preferred headlines and images that contained elements of 'A list' (17.60%), 'Game' (14.40%), 'Celebrity' (11.20%) and 'Quizzes' (10.40%).

| Coding labels | M (Section 1) | M (Section 3) | F (Section 1) | F (Section 3) |
|---|---|---|---|---|
| | Percentage | | | |
| Food | 0.00% | 0.00% | 4.32% | 4.80% |
| Accident news | 0.00% | 0.00% | 0.72% | 0.00% |
| Animals | 0.00% | 0.00% | 2.88% | 3.20% |
| Make up | 0.00% | 0.00% | 2.88% | 3.20% |

| | | | | |
|---|---|---|---|---|
| Dress | 0.00% | 0.00% | 2.16% | 2.40% |
| Social media news | 0.00% | 0.00% | 0.72% | 0.80% |
| Education | 0.00% | 0.00% | 0.72% | 0.80% |
| Advertisement | 0.00% | 0.00% | 0.72% | 0.80% |
| Shopping | 1.43% | 1.69% | 1.44% | 1.60% |
| Life sharing story | 1.43% | 0.00% | 2.16% | 2.40% |
| Health | 2.86% | 3.39% | 5.04% | 5.60% |
| A new or unknown concept | 5.71% | 6.78% | 2.88% | 2.40% |
| Travel | 7.14% | 8.47% | 0.72% | 0.80% |
| Quizzes | 7.14% | 8.47% | 9.35% | 10.40 % |
| You or I | 10.00% | 11.86% | 8.63% | 9.60% |
| Celebrity | 11.43% | 10.17% | 10.07% | 11.20 % |
| Political news | 15.71% | 5.08% | 15.83% | 8.00% |
| A list | 17.14% | 20.34% | 15.83% | 17.60 % |
| Game | 20.00% | 23.73% | 12.95% | 14.40 % |

**Table 9. Gender group click preference comparison between Section 1 and Section 3**

For different age groups, similarly, the main difference was the significant decrease in the percentage of 'Political news' for both groups. The percentage of 'Political news' headlines decreased from 10.50% to 6.43% for the young group and from 50% to 15.38% for the old group. After removing the non-clickbait data, it was shown that the young group preferred headlines and images that contained elements of 'Game' (18.71%), 'A list' (17.54%), 'Celebrity' (11.11%) and 'You or I' (10.53%). Meanwhile, the old group preferred headlines and images that contained elements of 'A list' (30.77%), 'Travel' (23.08%) and 'Political news' (15.38%).

| **Coding** | **Young group (Section 1)** | **Young group (Section 3)** | **Old group (Section 1)** | **Old group (Section 3)** |
|---|---|---|---|---|
| | **Percentage** | | | |

| | | | | |
|---|---|---|---|---|
| Game | 17.68% | 18.71% | 0.00% | 0.00% |
| Food | 3.31% | 3.51% | 0.00% | 0.00% |
| Accident news | 0.55% | 0.00% | 0.00% | 0.00% |
| Animals | 2.21% | 2.34% | 0.00% | 0.00% |
| Make up | 2.21% | 2.34% | 0.00% | 0.00% |
| Shopping | 1.66% | 1.75% | 0.00% | 0.00% |
| Dress | 1.66% | 1.75% | 0.00% | 0.00% |
| Social media news | 0.55% | 0.58% | 0.00% | 0.00% |
| Quizzes | 9.94% | 10.53% | 0.00% | 0.00% |
| Life sharing story | 1.66% | 1.75% | 3.57% | 0.00% |
| Education | 0.55% | 0.58% | 0.00% | 0.00% |
| Advertisement | 0.55% | 0.58% | 0.00% | 0.00% |
| Health | 4.42% | 4.68% | 3.57% | 7.69% |
| Celebrity | 10.50% | 11.11% | 10.71% | 7.69% |
| A new or unknown concept | 3.87% | 3.51% | 3.57% | 7.69% |
| You or I | 9.94% | 10.53% | 3.57% | 7.69% |
| Political news | 10.50% | 6.43% | 50.00% | 15.38 % |
| Travel | 1.66% | 1.75% | 10.71% | 23.08 % |
| A list | 16.57% | 17.54% | 14.29% | 30.77 % |

**Table 10. Gender group click preference comparison between Section 1 and Section 3**

# CHAPTER 5

# Discussion

## 5.1. Summary of results

In order to discover people's reactions to clickbait and their demographic trends, the results were based on three sections: 1. Overall analysis based on all the headlines or images that were clicked by the participants; 2. Detect clickbait headlines or images; 3. Results based on clickbait headlines, including comparison with section 1.

Section 1 showed that 'A list,' 'Political news,' 'Game' and 'Celebrity' were the four most attractive elements within all the headlines. Similarly, these four elements were the most attractive elements for both genders, although males had a centralised preference distribution of only 10 coding elements but females had a decentralised preference distribution of all 19 coding elements. The headlines that were clicked by males did not contain any elements of 'Food,' 'Accident news,' 'Animals,' 'Makeup,' 'Dress,' 'Social media news,' 'Education' and 'Advertisement.' Meanwhile, the young group preferred headlines that contained the same four elements with a decentralized distribution of 19 coding elements. However, the old group were not interested in game the 'Game' element but had an obvious preference for the 'Political news' and 'Travel' elements.

Section 2 showed that only 15.89% of clicked headlines were non-clickbait headlines, 83.44% of headlines were normal clickbait headlines and 0.66% of headlines were pure clickbait (the headline contained advertising or the link led directly to advertisements). Only one participant clicked on one advertising clickbait. It indicates that people can be enticed to clink on 'normal' clickbait headlines or images, however, they were able to identify and ignore pure clickbait headlines which contained advertising or promotional information during the experiment. In other words, participants were conscious of pure clickbait, however, they did not avoid 'normal' clickbait headlines. Moreover, the results showed that 70.80% of non-clickbait headlines contained only one coding elements and the remaining 29.20% contained two coding elements. In contrast, almost 60% of clickbait headlines contained two or more coding elements. Compared with non-clickbait headlines, clickbait headlines chose to cover more popular elements in order to attract people's attention.

In section 3, after removing the non-clickbait data from section 1, the results showed that the 'Political news' element was no longer the most attractive element as the majority of non-clickbait headlines were about 'Political news.' 'A list,' 'Game' and 'Celebrity' were the three most attractive elements for both genders. Apart from those elements, males were interested in 'You or I' elements and females were more interested in

‘Quizzes.’ Both age groups were attracted by ‘A list’ style clickbait headlines. However, unlike the young group who preferred ‘Game’ and ‘Celebrity’ elements, the old group preferred clickbait headlines that contained ‘Travel’ and ‘Political news’ elements.

During this experiment, the author also collected data on the participants’ opinions on the headline and image (Which one is more important? Which one motivates them to decide to click? Headline or Image). The data was collected by asking questions during the replay sessions and the audio recordings were saved using the University's drop-off service. The results showed that headline was more important than the image because although images can attract people’s attention, the click decisions were mainly based on reading the headlines. Furthermore, the majority of participants lost track of time during this experiment. Unlike pure advertising clickbait, people are willing to spend time reading the content of normal clickbait headlines. This finding matched the business model of some online media websites who produce readable and shareable content to take advantage of many different platforms to collect a large volume of readers and raise money by selling advertising space on their website.

It should be noted that while ‘Game’ was proven to be one of the most attractive elements used by clickbait headlines, all of the ‘Game’

elements that were clicked by participants in this experiment were about a specific game; ‘Pokémon go.’ Therefore, the excellent performance of the ‘Game’ element may only apply for phenomenal games or games with trending topics.

### 5.2. General discussion on clickbait

Generally speaking, clickbait is a pejorative word which refers to tricks and advertisements. However, there is no denying that clickbait and clickbait publishers are adjusting and developing along with the development of knowledge and society. In the past, the majority of clickbait links led directly to advertisements because the user would be taken to several different advertising websites after clicking on the clickbait links. These links usually attract people with exaggerated headlines and enticing pictures. For example, the headline can be ’10 best ways to earn £1 million in a month’ with a picture of a successful business man or a lot of cash. These links are straightforward and simple. They take advantage of people’s desire and curiosity. However, they are easy to avoid and once the user realises that it is a trick they will not be deceived again. Moreover, from the advertisers’ point of view, although it increased the click through counts, the effect of the advertisement is not that good because users can become angry at being tricked and close the advertising websites immediately. Only referral websites as an advertising vehicle can

earn money through online advertising business models such as pay-per-click.

Furthermore, some of the clickbait links are fraud and aim to infect the visitor's computer with a Trojan virus for stealing money. Therefore, the old fashioned clickbait links are often associated with crime risk.

However, clickbait links are becoming smarter and more difficult to define. They still use sensational headlines to attract people's attention but they also provide content that is related to the headline. The definition of clickbait seems infelicitous because the advertisers are becoming implicit. The adverts can be designed within a list of story units on the website and website owners can hide brands inside the headline content, blurring the demarcation between content and advertising.

Nowadays, clickbait links with direct advertising such as '10 best ways to earn £1 million in a month' are on the wane. The new trend of clickbait is to raise money through the user scale effect. This means that websites are aiming to build a large audience with high user viscosity. The headlines are designed and separated into detailed segments and follow the trending topics of political news and entertainment. In this stage, these websites, such as Buzzfeed, take advantage of social networking platforms like Facebook, Twitter and Snapchat, to make a chain of sharing-clicking-sharing to collect a large amount of human traffic. With a large user group,

Buzzfeed can sell side-screen banners or pop-up advertising at a higher price and those adverts will travel with the clickbait links through different social network platforms.

According to our experimental results, the majority of participants lost track of time while browsing the Buzzfeed website. This suggests that clickbait content is becoming more readable. Unlike the 'pure trick clickbait,' people are willing to spend time on the new forms of clickbait content, which will definitely increase the advertisers' evaluation on referral websites and increase the advertising cost. Furthermore, the Editor-in Chief of Buzzfeed, Ben Smith wrote an article called 'Why Buzzfeed Doesn't Do Clickbait.' Although this statement was not accepted by users, we can still observe the trend of clickbait development: the definition of clickbait and the distinction between normal headlines and clickbait is blurred.

The clickbait headlines are changing and they are not going to disappear because the online advertising business model motivates editors to create clickbait in order to attract readers. In fact, although clickbait links are becoming readable and shareable, it is not a positive trend for the media industry. This is because the quantity of information is much more important than the quality of the content in online media websites like Buzzfeed. They aim to produce as many articles as possible which are on

trend and can be shared rapidly. However, the quality, accuracy and objectivity cannot be guaranteed. As these kinds of websites are becoming a main channel for the young generation to accept information, the negative effects of those untraditional media companies need to be considered.

### 5.3. Eye tracking, participants' behaviours and the Hawthorne effect

In this study, the participants were arranged in the lab room and their eye movements were recorded by an eye tracker. Unlike Morae, which tracks and records the user's interactions with a website, the SME eye tracker can records the user's mouse movements and click activities, and people's eye movements and fixation time on different targets. The combined functionalities provide a better display of human performance and help us observe the participants' reactions to clickbait. We used the heat map tools to replay the participants' eye movements. The positions that the participants watched were colour coded. A longer fixation was marked in red and the shorter fixation was marked in yellow and then green. The colours were changed along with the eye movements of the participants. Eye tracking is interesting and attractive to the participants, however, the ethical issues should be considered. Eye movements are largely involuntary and may portray subliminal behaviour. Hence, the

participants can quit at any time and they can refuse to answer questions during the replay session.

The eye tracking method was designed to decrease or eliminate the bias caused by the Hawthorne effect because eye movements are more difficult to control and falsify than click activities. For example, if a participant spends longer on a headline or a sexy picture without clicking on it then it is possible that the participant wants to hide his true reaction. The inconsistency of eye movements and click activities can be easily observed, enabling us the ask questions to avoid bias.

However, in reality, the circumstances and the participants' behaviours were different from our original guess and the research became more complex. First, during the experiment, it was rare to see sexy pictures. Therefore, we did not observe the circumstance in which people looked at a sexy picture but chose not to click on it. This might be because of Buzzfeed's new trend towards being more legal and healthy. Instead it was observed that it was common for participants to look at (non-sexy) content for long but not click on it.

However, when asked, a typical response was that they knew the whole story of the link after reading the headline carefully. In that case, it was unnecessary to click on the headline because the user's curiosity had been satisfied by reading the headline. This is why some clickbait headlines

use shock elements and hide part of the story in order to keep the reader's curiosity.

Another reason of the inconsistency between eye movements and click activities was the language gap. A person whose first language is not English may not understand some English allusions or slang in some headlines because of their different cultural background. Therefore, the longer time spent on those headlines does not mean that the reader is interested in them, giving rise to an inconsistency. Moreover, some participants spent longer on a title because they were attracted by the picture and then determined that it was uninteresting after reading the headline. In addition, sometimes the participant intended to click and then realised that it was clickbait.

As a lab-based experiment, we cannot reproduce the same conditions as real life and the Hawthorne effect cannot be completely eliminated. However, we took steps to decrease the bias as much as possible. First, the participants conducted the experiment in a private room. They could browse the website naturally and freely without any disturbance. Second, they were clearly informed that their data would be anonymous and would only be used for their study. Their identifying information would be securely stored in accordance with the University's regulations. Third, using the eye tracker helped us observe people's interactions more clearly

and in greater detail, therefore, the results were closer to the real life experience.

### 5.4. Limitations

As we used an eye tracker in this experiment, one challenge is time. First, because the equipment is very expensive and we can only borrow it from the University of York's Computer Science department, the borrowing time is limited by the department's regulations. The maximum borrowing time is two hours a day, from Monday to Friday. This constraint meant that it was difficult to use a large sample in this study. This also applied to the analysis software as it was part of the equipment and could not be downloaded elsewhere due to the cost and copyright issues. Finally, the borrowed machine can only be used inside the Computer Science Building, therefore, the experiment was conducted in the Usability Lab.

The participants had to spend time travelling to the Heslington East campus and some participants who lived on the Heslington west campus or off campus refused to join this experiment because of the time problem, although they were interested in this experiment. The author of this study had follow the two hour rule when booking the lab and be considerate of the other researchers and students who were using these facilities.

Another limitation of this study is the inefficient data collection process. In this study, one of the most important aspects was to record the

clickbait headlines that the participants clicked on. Although the eye tracker machine would record all of the interactions between the participants and website, the headlines had to be manually typed when transcribing the session. However, in the replaying section, the main point was observe the performances of the participants and ask them questions about their click decisions, therefore, the headlines had to be recorded separately. This meant that the 15-20 minute eye tracker recording needed to be replayed at least three times to collect all of the data.

One of the original purposes of this study was to discover if different age groups have different reactions to clickbait headlines. However, only four participants were over 35 years old (including the pilot study participant). One reason is that the constraints on this study required people to come to the campus building, which is especially inconvenient for old people. As in other student studies, the author do not have a sufficient budget to invite and motivate old people to participate. Furthermore, these four participants were teachers or researchers working in the Computer Science department. They are more familiar with the internet and computers and their professional knowledge of this field can help them avoid clicking on clickbait headlines. Therefore, the results of old people's reactions may not be representative of the entire age group.

This study recruited participants who are from different countries, including Great Britain, Australia, China, Switzerland, Brazil and Saudi Arabia and the different cultural backgrounds may influence the result. Although Buzzfeed is an international media website which covers news and entertainment information from all over the world, the design of the headlines and contents mainly targets English native speakers. Therefore, a participant whose first language is not English may react differently to English clickbait. This effect was decreased because all the participants have higher education backgrounds and use English frequently. Furthermore, during the replay session, the author confirmed with them if they were interested in a headline or just attempting to understand it because of the language gap when they spent longer on a link. However, this aspect is a limitation of this experiment and should be considered in future research.

In addition, this experiment did not include the aspect of sharing clickbait via social networks. In daily life, apart from browsing online media websites, many people are attracted by clickbait which they then share on Facebook, Twitter or Snapchat. As mentioned in the literature review, Buzzfeed is using a different strategy to raise money, which is producing shareable content to take advantage of many different platforms. However, in a lab-based experiment, it is difficult to require the use of

social networks and observe the corresponding behaviours. It would be an ineffective method of discovering their responses the clickbait and an unethical invasion of their privacy. Moreover, the participants might be more inhibited in their responses if they were required to use their private social network account. Therefore, in this study, the experiment did not focus on clickbait sharing on social network media but concentrated on a website containing a lot of clickbait. However, the relationship between clickbait and social networks is very interesting and valuable to be considered in future research.

# CHAPTER 6

## Conclusion

This paper comprehensively reviews the relevant published sources about clickbait, regarding forward-referring, curiosity gap, social network media, eye movements and clickbait business model, which gives readers a clear view of clickbait's working principles. The study presents the eye tracking approach to observe people's reaction to clickbait and uses the qualitative research method for data analysis. The goal is to discover whether people react to clickbait and which kinds they are attracted to, with consideration of demographic trends. Consequently, we found that: 1. Generally, participants are attracted by clickbait headlines that contained elements of 'a list,' 'game,' 'celebrity' and 'You or I;' 2. Males' interests are centralised, focusing on 'a list,' 'game' and 'celebrity.' These were also the three most attractive elements for females, however, their interests are decentralised, covering all 19 coding elements; 3. The old group's preferences are centralised, focusing on 'a list,' 'travel' and 'political news.' In contrast, young people prefer 'a list,' 'game' and 'celebrity' and their preference distribution was decentralised; 4. People are immersed in the reading content after clicking on a clickbait headline, however, they avoided pure clickbait. Furthermore, the paper presents a discussion of clickbait, eye tracking and the limitations of this study.

The discussion helps to generate ideas for further work. Based on this study, people can design a more focused experiment for different groups with consideration of the relevant coding elements. The groups can be subdivided. For example, people who speak different languages should be divided into different groups and tested separately. Do different cultures and languages influence the design and performance of clickbait? In addition, researchers should note the importance of the connection between clickbait and social networking platforms and design appropriate experiment. Specifically, they should investigate which elements encourage people to share clickbait on their social network platforms. In addition, future research should consider clickbait on other devices, such as mobile phones, as people often use mobile phones to log in to their social network accounts and their reactions to clickbait may differ when using mobile applications.

# Reference

[1] Schaffer, Deborah, “Shocking secrets revealed! The language of tabloid headlines.” ETC: Rev. Gen. Semant. 1995. 52 (1), 27--46.

[2] Molek-Kozakowska, Katarzyna, “Towards a pragma-linguistic framework of the study of sensationalism in news headlines.” Discourse Commun. 2013. 7 (2), 173--197.

[3] Wallberg, Filip. “I Like News -- nyheder på Facebook [I like News -- news on Facebook].” 2013. ePub.

[4] Lindemann, Bernard, “Cheap thrills we live by: some notes on the poetics of tabloid headlinese.” J. Lit. Semant. 1990. 19 (1), 46--59.

[5] Y. Chen, N. J. Conroy, and V. L.Rubin, "Misleading Online Content: Recognizing Clickbait as "False News"," in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, Seattle, WA, 2015. pp. 15-19. Cho, J., Boyle, M. P., Keum.

[6] B. Adler, (2014). Tabloids in the age of social media. Columbia Journalism Review. [Online]. Available: http://www.cjr.org/news_literacy/national_enquirer_hoffman__hoax.php

[7] J. N. Blom & K. R. Hansen, “Click bait: Forward reference as lure in online news headlines.” In *Journal of Pragmatics*, 2015. 76, 87-100.

[8] B. Dervin, “From the mind's eye of the user: the sense-making qualitative and quantitative methodology.” *In J. D. Glazier, & R. R. Powell, Qualitative research in information management* (pp. 61-84). 1992. Englewood, CO: Libraries Unlimited.

[9] G. Loewenstein, “The psychology of curiosity: A review and reinterpretation.” In *Psychological Bulletin*, Vol 116(1), Jul 1994, 75-98.

[10] B. Frampton (2015, Sep 10). Clickbait: The changing face of online journalism

[Online]. Available: http://www.bbc.com/news/uk-wales-34213693

[11] Oxford Dictionaries. Clickbait [Online]. Available: https://en.oxforddictionaries.com/definition/clickbait

[12] S. Pengnate (2016). Measuring Emotional Arousal in Clickbait: Eye-Tracking Approach [Online]. Available: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1133&context=amcis2016

[13] M. Potthast, S. K¨opsel, B. Stein, and M. Hagen, "Clickbait Detection," in Advances in Information Retrieval, 38th European Conference on IR Research, ECIR 2016, Padua, Italy, March 20-23, 2016. Proceedings, pp. 810–817.

[14] C. Jenkins. (2014, Feb 15) "Kfc: The Story of a 3-Year-Old Girl Being Asked to Leave Was a Hoax." [Online]. Available: http://www.businessinsider.com/scarred-mississippi-girl-asked-leave-kfc-untrue-2014-6

[15] G. Sandoval. (2016, Feb 15) "Who's to Blame for Spreading Phony Jobs Story?" [Online]. Available: www.cnet.com/news/whos-to-blame-for-spreading-phony-jobs-story/

[16] S. M. Kim, A.Valitutti, and R. A. Calvo. "Evaluation of Unsupervised Emotion Models to Textual Affect Recognition," in Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, CA, pp. 62-70.

[17] J. A. Gibbons, A. F. Lukowski and W. R. Walker. "Exposure Increases the Believability of Unbelievable News Headlines Via Elaborate Cognitive Processing," Media Psychology (7:3), 2005. pp. 273-300.

[18] T. Doyle. (2016, Feb 19) "Don't Click That: The Clickbait Virus." [Online]. Available: http://www.ahundredmonkeys.com/clickbaitadvertising/

[19] E. C. Tandoc. "Journalism Is Twerking? How Web Analytics Is Changing the

Process of Gatekeeping," New Media & Society (16:4), 2014. pp. 559-575.

[20] J.N. Blom, K.R. Hansen. “Click bait: forward-reference as lure in online news headlines.” J. Pragmat. 76, 87–100 (2015)

[21] Ifantidou, Elly, Newspaper headlines and relevance: ad hoc concepts in ad hoc contexts. J. Pragmat, 2009. 41, 699--720.

[22] J. Dvorkin, “Column: Why click-bait will be the death of journalism,” pbs.org/newshour/making-sense/what-you-dont-know-aboutclick-bait-journalism-could-kill-you/.

[23] G. Mark, “Click bait is a distracting affront to our focus,” nytimes.com/roomfordebate/2014/11/24/you-wont-believe-whatthese-people-say-about-click-bait/click-bait-is-a-distracting-affrontto-our-focus.

[24] P. Koechley. (2012) Why the title matters more than the talk. [Online]. Available: http://blog.upworthy.com/post/26345634089/why-the-title-matters-more-than-the-talk

[25] K. El-Arini, J. Tang. (2014) News feed FYI: click-baiting. [Online]. Available: http://newsroom.fb.com/news/2014/08/news-feed-fyi-click-baiting/

[26] A. Chakraborty, B. Paranjape, S. Kakarla and N. Ganguly. (2016) Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media. [Online]. Available: http://www.facweb.iitkgp.ernet.in/~niloy/PAPER/P-ASONAM-2016_paper_64.pdf

[27] D. Salvucci, J. Goodberg, Identifying fixations and saccades in eye-tracking protocols. Proceeding ETRA '00 Proceedings of the 2000 symposium on Eye tracking research & applications Pages 71-78. [Online]. Available: http://dl.acm.org/citation.cfm?id=355028&CFID=669794998&CFTOKEN=10673243

[28] K.Rayne. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124: 372--422, 1998.

[29] M. M. Bradley, L.Miccoli, M. A Escrig and P. J. Lang, "The Pupil as a Measure

of Emotional Arousal and Autonomic Activation," Psychophysiology (45:4), 2008. pp. 602-607.

[30] R. F. Stanners, M. Coulter, A. W. Sweet and P. Murphy, "The Pupillary Response as an Indicator of Arousal and Cognition," Motivation and Emotion (3:4), 1979. pp. 319-340.

[31] T. Partala, M. Jokiniemi and V. Surakka. "Pupillary Responses to Emotionally Provocative Stimuli," in Proceedings of the 2000 Symposium on Eye Tracking Research & Applications, Palm Beach Gardens, FL, pp. 123-129.

[32] P. Desmet. "Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Products," in Funology: From Usability to Enjoyment, M.A. Blythe, K. Overbeeke, A.F. Monk and P.C. Wright (eds.). Dordrecht: Springer Netherlands, 2005. pp. 111-123.

[33] C. M. Masciocchi, and J. D. Still. "Alternatives to Eye Tracking for Predicting Stimulus-Driven Attentional Selection within Interfaces," Human-Computer Interaction (28:5), 2013. pp. 417-441.

[34] Myllylahti, 2014 Newspaper paywalls – the hype and the reality Digital Journal., 2 (2) (2014), pp. 179–194

[35] Thurman and Herbert, 2007 Neil Thurman, Jack Herbert Paid content strategies for news websites Journal. Pract., 1 (2) (2007), pp. 208–226.

[36] Greer, Advertising on traditional media sites: can the traditional business model be translated to the web? Soc. Sci. J., 41 (1) (2004), pp. 107–113.

[37] V. Anupama, A. Mayera, K. Nissimb, B. Pinkasb and M. Reitera, “On the security of pay-per-click and other Web advertising schemes” in *Computer Networks,* Volume 31, Issues 11–16, 17 May 1999, Pages 1091–1100.

[38] M.K. Reiter, V. Anupam and A. Mayer, Detecting hit shaving in click-through

payment schemes, in: Proc. 3rd USENIX Workshop on Electronic Commerce, September 1998, pp. 155–166.

[39] K. L. Hartog, System and method for pay-per-click revenue sharing (2006). [Online]. Available: https://www.google.com/patents/US20060080239#forward-citations

[40] H. Guinness, I Followed Click Bait Internet Ads, Here's What I Found (2015). [Online]. Available: http://www.makeuseof.com/tag/followed-click-bait-internet-ads-heres-found/

[41] A. M. Spence, Notes on Advertising, "Economies of Scale, and Entry Barriers" in *The Quarterly Journal of Economics* Vol. 95, No. 3 (Nov., 1980), pp. 493-507.

[42] E. Dash. (2016, Apr 25). Buzzfeed's Untraditional Business Model Spurs Financials Inquiry [Online]. Available: http://theticker.org/buzzfeeds-untraditional-business-model-spurs-financials-inquiry

[43] H. McKenzie. (2013, Jan 9). How does BuzzFeed intend to crack a $1B valuation? By embracing journalism's past [Online]. Available: https://pando.com/2013/01/09/how-does-buzzfeed-intend-to-crack-a-1b-valuation-by-embracing-journalisms-past/

[44] H. Blodget. (2013, Jan 9). BUSINESS INSIDER SECRETS REVEALED [Online]. Available: http://www.businessinsider.com/business-insider-traffic-2013-1?IR=T#

[45] D. Frommer. (2010, Jul 6). Here's How BuzzFeed Plans To Turn Viral Content Into Cash [Online]. Available: http://www.businessinsider.com/heres-how-buzzfeed-plans-to-turn-viral-content-into-cash-2010-7?IR=T

[46] Buzzfeed. (2015, Aug 22). About BuzzFeed [Online]. Available: https://www.buzzfeed.com/about

[47] Buzzfeed Video on YouTube. (2014, Oct 24) [Online]. Available: https://www.youtube.com/user/buzzfeedvideo/about

[48] B. Smith. (2015, Feb 27). The Dress [Online]. Available: https://www.buzzfeed.com/bensmith/culture-web-culture?utm_term=.jv399Wz27#.nyxNN0eOY

[49] C. Liam. (2014, Oct 23). The Biggest Facebook Publishers of March 2014 [Online]. Available: https://www.newswhip.com/2014/04/biggest-facebook-publishers-march-2014/

[50] Pew Research Center. (2014, Oct 21). Where News Audiences Fit on the Political Spectrum [Online]. Available: http://www.journalism.org/interactives/media-polarization/table/overall/

[51] M. Sweney. (2016, Jan 13). BuzzFeed breaks UK ad rules over misleading advertorial [Online]. Available: https://www.theguardian.com/media/2016/jan/13/buzzfeed-breaks-uk-ad-rules-over-misleading-advertorial

[52] R. McCarney, J. Warner, S. Iliffe, R. Haselen, M. Griffin, P. Fisher. "The Hawthorne Effect: a randomised, controlled trial". BMC Med Res Methodol. 2007. 7:30.doi:10.1186/1471-2288-7-30. PMC 1936999 . PMID 17608932.

[53] H. A. Landsberger, *Hawthorne Revisited*, Ithaca, 1958.

[54] D. M. McBride. The process of research in psychology. London: Sage Publications. 2013.

# Appendix A: Consent Form

Participant Identification Number:

## CONSENT FORM

**PROJECT TITLE:** A study on browsing experience
Name of Researcher: Yiteng Xing
Project Supervisor: Dr Alistair Edwards, alistair.edwards@york.ac.uk

**Please initial box**

1. I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐

3. I understand that any information given by me may be used in future reports, articles or presentations by the research team. ☐

4. I understand that I will not be identifiable in any reports, articles or presentations. ☐

5. I agree that an audio recording can be made of my interaction with the researcher. ☐

I agree to take part in the above study.

_____ _____ _____

Name of Participant Date Signature

Yiteng Xing

| | | |
|---|---|---|
| _____ | _____ | |
| | _____ | |
| Researcher | Date | Signature |

# Appendix B: Information Sheet

## A STUDY ON BROWSING EXPERIENCE: INFORMATION SHEET

**PROJECT TITLE**

A study on browsing experience

**INVITATION**

You are being asked to take part in a research study on the project about analysing people's browsing experience. This project is being conducted by Yiteng Xing as the Individual Project Module of the master's programme in Human-Centred Interactive Technologies in the Department of Computer Science at the University of York. This project is supervised by Dr Alistair Edwards who has overall responsibility for the conduct of this study under the approval of the Physical Sciences Ethics Committee at the University.

**WHAT WILL HAPPEN**

In this study you are asked to imagine that you have 15 minutes to spare and you have decided to pass the time surfing the web. You will be shown a web page and it will be up to you what you want to view on that page and whatever links you might follow.

While you are doing this your interaction with the browser (i.e. movements of the cursor, clicks and any typing) will be recorded. Also your eye movements (i.e. where you are looking on the screen) will be tracked and please do not move around during that time.

You will be left on your own to browse. When you think 15 minutes is up you can stop. If you have not finished after 20 minutes, though, I will come and stop you. Please try not to look at the watch. (We are interested as to whether you lose track of time during the activity).

You will then be given further, debriefing information. We will also collect some basic information about yourself. Then you will be shown the recordings of your session, during which you will be asked to explain some of your behaviour. For instance, you might be asked, 'Why did you decide to click on that particular link?' An audio recording will be made of this part of the session.

**TIME COMMITMENT**

The study typically takes 50 minutes in total.

**PARTICIPANTS' RIGHTS**

You may decide to stop being a part of the research study at any time up to the end of today's session without explanation. You have the right to ask that any data you have supplied to that point be withdrawn and destroyed.

You have the right to omit or refuse to answer or respond to any question that is asked of you.

You have the right to have your questions about the procedures answered. If you have any questions as a result of reading this information sheet, you should ask the researcher before the study begins.

As you click through the web you may encounter material that might be considered offensive. You are not obliged to view any such materials and can click back to the previous page. You may also withdraw from the study if you do encounter anything offensive.

**BENEFITS AND RISKS**

There are no known benefits or risks for you in this study.

**COST, REIMBURSEMENT AND COMPENSATION**

Your participation in this study is voluntary. However, all participants have the option of being entered into a draw from which two winners will be chosen at random to receive a £20 Amazon voucher.

**CONFIDENTIALITY/ANONYMITY**

You will be allocated a Participant Number. That will be used to identify your data, and the link between that number and your identity will be stored securely.

The experiment will be written up as a report for my assessment on the module. It may also be published in academic outlets such as journals, conferences or research books. In all cases, the data will only be presented in summary form and you will not be directly identifiable in any way.

**FOR FURTHER INFORMATION**

Alistair Edwards and I will be glad to answer your questions about this study at any time. You may contact him at alistair.edwards@york.ac.uk, or contact me at yx1058@york.ac.uk.

If you want to find out about the final results of this study, please let me know now and I will email you the results upon completion of the project.

# Appendix C: Debriefing sheet

Participant Identification Number: Time Starts at:

Time Ends at:

Duration:

## Debriefing sheet

[Note that the debriefing will take place *after* the rest of the session is complete – including the replay of the web browsing].

Thank you so much for participating in this study. Your participation was very valuable to us. We know you are very busy and very much appreciate the time you devoted to participating in this study.

### Demographic information

Please answer the questions below. You are at liberty not to answer any of them if you prefer.

1. Name

2. Sex:
   - Male ☐
   - Female ☐
   - Prefer not to say ☐

3. Age

4. What is your occupation? If student, then what are you studying and at what level (e.g. MSc in Human-Centred Interactive Technologies).

5. What is the highest level of educational qualification you have (not including the programme you are currently on, if you are a student).
   - Pre-secondary school (e.g. GCSE) ☐
   - Secondary school (e.g. A-Level) ☐
   - Bachelor's degree ☐
   - Master's degree ☐
   - PhD ☐
   - Other (please explain) ☐

6. What is your first language?

7. Roughly how much time do you spend surfing the web? That is to say that you might be doing multiple tasks on a computer, tablet or phone, but you generally have a browser open.
   - Several hours per day (i.e. more than 6 hours on and off). ☐
   - A few hours per day (less than 6 hours). ☐
   - Approximately one how per day. ☐
   - A few hours per week ☐
   - Very occasionally. ☐
   - I don't really do any web surfing. ☐

8. Why do you surf the web? Please estimate the percentage of time you spend on these purposes. For instance, if you use the web mainly for work, but occasionally glance at Twitter, you might allocate 90% to *Work/study,* 10% to *Entertainment* and 0% to *Other.*
   - Work/study ___ %
   - Entertainment ___ %
   - Other ___ %
   - (Please describe)

9. Please list any social media sites that you visit regularly (e.g. Facebook, Twitter, Instagram,…)

   _____
   _____
   _____
   _____

10. Do you have any hobbies? Please list the **three** main activities which you undertake when you are not working.
    1. _____
    2. _____
    3. _____

11. Have you ever bought anything as a result of seeing an advert on the Web? That is to say that you were not on the web with the intention of buying something, but you saw an advert for something and went on to buy it?
    - Yes, frequently ☐
    - Yes, once or twice ☐
    - I don't think so ☐
    - No. ☐

## Debriefing

This study is about *Clickbait*. 'Clickbait' refers to websites which have insubstantial content, but a lot of advertising. It also refers to links to such sites: the 'bait' which attracts visitors to the site.

You probably noticed that the webpage I gave you at the start contained a lot of clickbait. We are interested to see which – if any – of the bait worked for you. We hope to be able to find out how the different types of clickbait work with different people

I was unable to tell you this at the outset because that would have drawn your attention to the clickbait. For the same reason I would ask that you do not tell any of your friends who may be taking part in this study that this its objective.

If you have any questions or concerns, you may contact me at yx1058@york.ac.uk.

Thank you so much for your participation!

Now that you have full information about the experiment, please sign below if you are happy for me to use the data collected. If you are not happy, then the data will be destroyed.

Signed

Date.

If you wish to be sent a copy of the project report on completion, please give an email address to which it can be sent.

Also if you wish to be entered in the draw for an Amazon voucher, please give an email address to which the voucher can be sent.

Your address will not be used for any other purpose.

Please tick as appropriate

☐ I wish to be entered in the draw for a £20 Amazon voucher.

☐ I wish to receive a copy of the project report.

Email: _____