A Twitter-based study on the reach of a smoking cessation organisation and the social meaning of smoking

Cornelia Afra Maria van Diepen

This thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth

January 2018
Abstract

Tobacco smoking is one of the leading causes of death worldwide and, while anti-tobacco movements have implemented many policies and regulations to prevent smoking uptake, every year approximately 207,000 teenagers in the UK start smoking. Much has been written on youth smoking initiation, but examining the perspectives of young people through their Twitter persona has thus far been under-researched.

The Twitter engagement of the smoking cessation organisation The Filter Wales is used as the basis of this study. This organisation is a leading youth-focused tobacco control initiative in Wales. Through a Twitter collection program, 4.8 million tweets from 2180 sample members were collected. Novel big data methods are combined with traditional methods such as content analysis and analysis of discourse to determine the reach of smoking cessation programs and provide a better understanding of the social meaning of smoking and its association with other health risks.

Results show that the tweets and Twitter profiles can illustrate social inequalities between sample members and that the Filter has reached their target audience. The results further illustrate that the majority of tweets concern tobacco smoking and that for young people tobacco and e-cigarettes relate to personal behaviour while marijuana and shisha are more common in a social context. Important for the Filter Wales, the knowledge of smoking (and unhealthy co-behaviours) is present, but for young people, the positive short-term effects of unhealthy behaviour outweigh the long-term gains of a healthy lifestyle.

This thesis extends the large body of work by approaching smoking from a data-driven perspective which has not been done to this extent before. This thesis demonstrates how tweets provide unadjusted perceptions of smoking and uniquely shows how these young people can be better understood in their smoking habit.
Declaration

Whilst registered as a candidate for the above degree; I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 74,139

Signed: ........................................ Date: 31 January 2018
# Table of Content

**ABSTRACT** .......................................................................................................................... I
**DECLARATION** .........................................................................................................................II
**LIST OF TABLES** ....................................................................................................................... VII
**LIST OF FIGURES** ..................................................................................................................... IX
**LIST OF ABBREVIATIONS** ........................................................................................................ XIII
**ACKNOWLEDGEMENTS** ............................................................................................................ XIV
**DISSEMINATION** ....................................................................................................................... XX

## CHAPTER 1. INTRODUCTION ................................................................................................. 1

1.1 **THE BACKGROUND** ....................................................................................................... 1

1.1.1 **The Smoking Endgame** ............................................................................................. 3

1.1.2 **Smoking in the United Kingdom** ............................................................................... 4

1.1.3 **Youth smoking in Wales** ............................................................................................ 8

1.2 **THE PROBLEM** ................................................................................................................ 8

1.3 **THIS STUDY’S JUSTIFICATION AND AIMS** .................................................................. 10

1.4 **THE STRUCTURE OF THE THESIS** ............................................................................ 11

## CHAPTER 2. LITERATURE REVIEW ON YOUTH SMOKING .................................................. 14

2.1 **SMOKING INITIATION** .................................................................................................... 15

2.1.1 **Favourable opinions** .................................................................................................. 16

2.1.2 **Peer influence** ........................................................................................................... 17

2.1.3 **Experimenting with smoking** .................................................................................... 18

2.2 **SOCIAL SMOKING** ....................................................................................................... 19

2.2.1 **Towards nicotine dependence** .................................................................................. 20

2.2.2 **Developing the smoker identity** ................................................................................ 21

2.3 **THE QUITTING STAGE** ................................................................................................ 22

2.3.1 **Relapsing** .................................................................................................................... 24

2.3.2 **The hardcore smokers and quitting** ......................................................................... 25

2.4 **NON-SMOKERS** ............................................................................................................. 26

2.5 **TOBACCO CO-BEHaviours** ......................................................................................... 27

2.5.1 **Tobacco and e-cigarettes** ........................................................................................... 27

2.5.2 **Tobacco and marijuana** ............................................................................................. 28

2.5.3 **Tobacco and shisha smoking** ..................................................................................... 29

2.6 **TOBACCO AND OTHER HEALTH-RELATED BEHAVIOURS** .................................. 30

2.6.1 **Health co-behaviours and smoking initiation** .......................................................... 32

2.6.2 **Health co-behaviours in social settings** .................................................................... 32

2.6.3 **Quitting smoking and health co-behaviours** .............................................................. 33

2.7 **CONCLUDING REMARKS** ............................................................................................ 34

## CHAPTER 3. YOUTH SMOKING REGULATIONS AND INTERVENTIONS ................................. 36

3.1 **TOBACCO CONTROL POLICIES** .................................................................................. 37

3.1.1 **Reducing tobacco sales** ............................................................................................. 37

3.1.2 **Smoke-free places** ..................................................................................................... 38

3.1.3 **Prohibiting tobacco advertisement** .......................................................................... 39

3.1.4 **Indirect marketing as a response** .............................................................................. 41

3.2 **TOBACCO CONTROL: HELPING PEOPLE TO QUIT** .................................................. 42

3.2.1 **Anti-smoking campaigns** .......................................................................................... 42

3.2.2 **Youth-based initiatives** ............................................................................................. 44

3.2.3 **Online awareness and counselling** ............................................................................ 47

3.3 **CHALLENGES OF EXISTING APPROACHES** ............................................................... 48

3.3.1 **Marginalisation through smoking legislation** ............................................................ 48

---
CHAPTER 4. THE RESEARCH USE OF TWITTER DATA .............................................60
4.1 WHAT IS TWITTER .........................................................................................60
4.1.1 What is a ‘tweet’ .......................................................................................61
4.1.2 Followers on Twitter ...............................................................................62
4.2 TWITTER AS A RESEARCH TOOL ..............................................................63
4.2.1 Examples of Twitter as an academic research tool .......................................63
4.2.2 Challenges of traditional methods compared to Twitter research ..................65
4.2.3 Challenges of using Twitter as a research tool .............................................67
4.3 ETHICAL CONSIDERATIONS FOR TWITTER RESEARCH .........................68
4.3.1 Following ethical guidelines ......................................................................69
4.4 CREATING THE TWITTER DATA FOR THE PRESENT STUDY .....................70
4.4.1 Twitter as part of the Filter campaign .......................................................71
4.4.2 Data harvesting of The Filter Wales Twitter .................................................71
4.4.3 The young peoples' data harvesting .........................................................74
4.4.4 Twitter profiles .........................................................................................76
4.4.5 Derived variables ......................................................................................78
4.5 DATA PREPARATION ....................................................................................78
4.5.1 The smoking-related tweets dataset .........................................................79
4.5.2 Smoking status .........................................................................................80
4.5.3 Smoking type, person, and action ..............................................................81
4.5.4 Geolocation ..............................................................................................82
4.5.5 Determining place of residence ...............................................................84
4.5.6 Health risk behaviour concepts ...............................................................87
4.6 UNDERSTANDING THE TWEETS .............................................................88
4.6.1 Sarcasm and other sentiments .................................................................88
4.6.2 Understanding retweets ............................................................................89
4.7 CONCLUDING REMARKS .........................................................................90

CHAPTER 5. METHODS OF ANALYSIS ................................................................91
5.1 DESCRIPTIVE METHODS ..........................................................................91
5.1.1 Locational depictions .............................................................................92
5.1.2 Quantitative content analysis .................................................................93
5.1.3 Temporal analysis ..................................................................................94
5.1.4 Sentiment Analysis ..............................................................................95
5.2 QUALITATIVE METHODS .........................................................................98
5.2.1 Qualitative content analysis ..................................................................98
5.2.2 Analysis of the discourse ...................................................................100
5.2.3 Linguistic analysis .................................................................................101
5.3 CONCLUDING REMARKS .......................................................................103

CHAPTER 6. DESCRIBING THE STUDY SAMPLE .............................................104
6.1 THE GENDER, AGE AND SMOKING STATUS OF THE SAMPLE ..................105
6.1.1 Age and gender ....................................................................................105
6.1.2 Twitter activity according to gender ......................................................106
6.1.3 Twitter activity according to age ............................................................109
6.1.4 Smoking status characteristics .............................................................110
6.1.5 Smoking status according to age and gender .........................................111
6.2 MAPPING THE SAMPLE .........................................................................112
6.2.1 Tweets across the UK .........................................................................114
6.2.2 Using the geolocation information to reveal likely place of residence .......117
CHAPTER 10. CONCLUSIONS.................................................................................................................. 201
10.1 SUMMARY OF FINDINGS................................................................................................................ 202
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.1</td>
<td>Classification of the smoking status</td>
<td>p.80</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Classification of the smoking product</td>
<td>p.80</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Classification of the smoking activity</td>
<td>p.81</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Classification of Health Behaviours</td>
<td>p.86</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Categories of the linguistics analysis described by the Analyzewords program</td>
<td>p.100</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Summary data on age and gender</td>
<td>p.104</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Twitter engagement across the sample</td>
<td>p.106</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Summary data on the number of tweets made by the sample</td>
<td>p.106</td>
</tr>
<tr>
<td>Table 6.4</td>
<td>Summary data on the number of followings</td>
<td>p.107</td>
</tr>
<tr>
<td>Table 6.5</td>
<td>Summary data on the number of followers</td>
<td>p.107</td>
</tr>
<tr>
<td>Table 6.6</td>
<td>Summary data on the number of likes</td>
<td>p.107</td>
</tr>
<tr>
<td>Table 6.7</td>
<td>Smoking status across the study sample</td>
<td>p.110</td>
</tr>
<tr>
<td>Table 6.8</td>
<td>Smoking status among men and women</td>
<td>p.110</td>
</tr>
<tr>
<td>Table 6.9</td>
<td>Distribution of smoking status per age group</td>
<td>p.111</td>
</tr>
<tr>
<td>Table 6.10</td>
<td>Neighbourhood descriptive statistics across quintiles of WIMD ranks in Wales</td>
<td>p.125</td>
</tr>
<tr>
<td>Table 6.11</td>
<td>Smoking status of the Twitter users per WIMD scores in Wales</td>
<td>p.126</td>
</tr>
<tr>
<td>Table 6.12</td>
<td>Neighbourhood descriptive statistics with Rural/Urban classification in Wales</td>
<td>p.127</td>
</tr>
<tr>
<td>Table 6.13</td>
<td>Smoking status of the Twitter users per Urban/Rural Divide in Wales</td>
<td>p.128</td>
</tr>
<tr>
<td>Table 7.1</td>
<td>Descriptive summary of tweets according to smoking product, action and who the action refers to.</td>
<td>p.132</td>
</tr>
<tr>
<td>Table 7.2</td>
<td>Descriptive summary of tobacco-related tweets according to gender</td>
<td>p.134</td>
</tr>
<tr>
<td>Table 7.3</td>
<td>Descriptive summary of marijuana-related tweets according to gender</td>
<td>p.135</td>
</tr>
<tr>
<td>Table 7.4</td>
<td>Descriptive summary of e-cigarette-related tweets according to gender</td>
<td>p.135</td>
</tr>
<tr>
<td>Table 7.5</td>
<td>Descriptive summary of shisha-related tweets according to gender</td>
<td>p.136</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Table 7.6</td>
<td>Descriptive table on the type of product, who is mentioned in the tweet and the average sentiment score.</td>
<td>p.147</td>
</tr>
<tr>
<td>Table 7.7</td>
<td>Descriptive table on the type of product, the smoking action and the average sentiment score.</td>
<td>p.148</td>
</tr>
<tr>
<td>Table 7.8</td>
<td>Descriptive table on the type of product and the average sentiment score according to gender.</td>
<td>p.149</td>
</tr>
<tr>
<td>Table 8.1</td>
<td>Prevalence of health behaviours in smoking-related tweets divided by gender</td>
<td>p.159</td>
</tr>
<tr>
<td>Table 8.2</td>
<td>Prevalence of health behaviours in smoking-related tweets divided by age group</td>
<td>p.161</td>
</tr>
</tbody>
</table>
List of Figures

| Figure 1.1 | The ‘smoking epidemic’ model for high-income countries time frames (Source: Thun et al., 2012). | p.2 |
| Figure 1.2 | Proportion of British adults (16+) who smoke (1974-2012) (adapted from General Lifestyle Survey, 2013) | p.5 |
| Figure 1.3 | Proportion of British adults (16+) who quit smoking (1974-2012) (adapted from General Lifestyle Survey, 2013) | p.6 |
| Figure 1.4 | Proportion of British adults (16+) who have never smoked (1974-2012) (adapted from General Lifestyle Survey, 2013) | p.6 |
| Figure 1.5 | Adult smoking prevalence for Wales and UK (source: General Lifestyle Survey, 2011) | p.7 |
| Figure 1.6 | Smoking prevalence at different ages of men and women in Wales in 2015 (adapted from the Welsh Health Survey, 2015 & HBSC, 2015) | p.8 |
| Figure 3.1 | Advertisement poster for Sampoerna mild cigarettes with the quote “I contain noise.” (source: www.shootinggalleryasia.com/photographer-photo/a-mild-go-ahead-iga-massardi/) | p.39 |
| Figure 3.2 | Screenshot of a famous vlogger reviewing Marlboro Red on YouTube (accessed on 28/09/2017) | p.41 |
| Figure 3.3 | Poster of the 2017 ‘World No Tobacco Day’ campaign (source: www.who.int/tobacco/wntd/ accessed on 28/09/2017) | p.42 |
| Figure 3.4 | A still for a Truth® commercial about the increased chance of cancer for pets if the owner smokes (source:www.thetruth.com/ accessed on 28/09/2017) | p.45 |
| Figure 3.5 | Screenshot of the website of ASH Wales (www.ashwales.org.uk accessed on 13/12/16) | p.52 |
| Figure 3.6 | Screenshot of the ashwales.org.uk/en/about/our-youth-project (accessed on 03/05/2017) | p.53 |
| Figure 4.1 | Diagram of the data collection process | p.69 |
| Figure 4.2 | Screenshot of the raw Twitter data in Excel derived from the Twitter Company | p.70 |
| Figure 4.3 | Screenshot of the raw Twitter data from the API program in Excel | p.72 |
| Figure 4.4 | Screenshot of a Twitter profile | p.74 |
| Figure 4.5 | The SQL codes for the smoking-related terms | p.78 |
Figure 4.6  Screenshot of the output of the ArcGIS join of the coordinates and LSOA layers in Excel  p.84
Figure 5.1  Screenshot of the output of the AnalyzeWords program  p.101
Figure 6.1  The age distribution of the study sample according to gender  p.105
Figure 6.2  Distribution of Twitter engagement per age group  p.109
Figure 6.3  Distribution of Twitter geolocation coordinates on a world map  p.112
Figure 6.4  Distribution of Twitter geolocation coordinates within the United Kingdom and Ireland  p.114
Figure 6.5  Distribution of Twitter geolocation coordinates within Wales with an underlay of LSOA boundaries  p.115
Figure 6.6  Distribution of sample members per LSOA in Wales  p.117
Figure 6.7  Overview map of specific areas for close-up Figures 6.7a-d  p.118
Figure 6.7a  Count of sample members per LSOAs in the Swansea area (South Wales)  p.119
Figure 6.7b  Count of sample members per LSOAs in the Cardiff area (South Wales)  p.119
Figure 6.7c  Count of sample members per LSOAs in the Aberystwyth area (west coast of Wales)  p.120
Figure 6.7d  Count of sample members per LSOAs in the North of Wales  p.120
Figure 6.8  Scatterplot of the number of smoking-related tweets per percentage of youth population per local authority in Wales  p.122
Figure 6.9  Adults who report smoking by deprivation quintile (adapted from the Welsh Health Survey, 2014)  p.124
Figure 6.10 Prevalence adult smoking by Welsh Index of Multiple deprivation quintiles (adapted from the Welsh Health Survey, 2014)  p.124
Figure 7.1  Number of smoking type tweets per age group with percentages within each age group in brackets  p.137
Figure 7.2  Number of smoking type tweets per age group with percentages within each age group in brackets  p.138
Figure 7.3  Timeline of the number of tweets per hours of the day according to smoking product referred to in the tweets  p.140
Figure 7.4  Tweet frequency by day of the week according to the smoking product referred to in the tweets  p.142
Figure 7.5  Monthly tweeting activity according to the smoking product referenced in the tweet  p.143
| Figure 7.6 | Timeline of the frequency of tobacco-related tweets per week in the years July 2013 to June 2016 | p.144 |
| Figure 7.7 | Timeline of the percentage of tobacco tweets per months of the year divided by smoking action | p.145 |
| Figure 7.8 | Time series of average sentiment score for the tobacco and e-cigarettes tweets per hours of the day | p.150 |
| Figure 7.9 | Time series of average sentiment score for the marijuana and shisha tweets per hours of the day | p.151 |
| Figure 7.10 | Time series of average sentiment score for tobacco and e-cigarette tweets per days of the week | p.152 |
| Figure 7.11 | Time series of average sentiment score for marijuana and shisha tweets per days of the week | p.153 |
| Figure 7.12 | Timeline of average sentiment score for the tobacco and e-cigarette tweets per months of the year | p.153 |
| Figure 7.13 | Timeline of average sentiment score for the marijuana and shisha tweets per months of the year | p.153 |
| Figure 8.1 | Prevalence of other health behaviour references in smoking-related tweets | p.158 |
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>ASH</td>
<td>Action for Smoking and Health</td>
</tr>
<tr>
<td>ASSIST</td>
<td>A Stop Smoking in Schools Trial</td>
</tr>
<tr>
<td>CO</td>
<td>Carbon monoxide</td>
</tr>
<tr>
<td>CSV</td>
<td>Comma-separated values</td>
</tr>
<tr>
<td>DCLG</td>
<td>Department for Communities and Local Government</td>
</tr>
<tr>
<td>DEFRA</td>
<td>Department for Environment, Food &amp; Rural Affairs</td>
</tr>
<tr>
<td>ESFA</td>
<td>European Smoking prevention Framework Approach</td>
</tr>
<tr>
<td>ESRI</td>
<td>Environmental Systems Research Institute</td>
</tr>
<tr>
<td>FCTC</td>
<td>Framework Convention on Tobacco Control</td>
</tr>
<tr>
<td>FP</td>
<td>First Person</td>
</tr>
<tr>
<td>GIF</td>
<td>Graphic Interchange Format</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>HBSC</td>
<td>Health Behaviours in School-aged Children survey</td>
</tr>
<tr>
<td>LA</td>
<td>Local Authority</td>
</tr>
<tr>
<td>LIWC</td>
<td>Linguistic Inquiry and Word Count</td>
</tr>
<tr>
<td>LLSS</td>
<td>Liverpool Longitudinal Smoking Study</td>
</tr>
<tr>
<td>LSOA</td>
<td>Lower Layer Super Output Area</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Services</td>
</tr>
<tr>
<td>ONS</td>
<td>Office for National Statistics</td>
</tr>
<tr>
<td>RUC</td>
<td>Rural and Urban Classification</td>
</tr>
<tr>
<td>RYO</td>
<td>Roll Your Own</td>
</tr>
<tr>
<td>SDD</td>
<td>Smoking, Drinking and Drug Use among Young People in England</td>
</tr>
<tr>
<td>SEP</td>
<td>Socioeconomic position</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>WIMD</td>
<td>Welsh Indices of Multiple Deprivation</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
Acknowledgements

Special thanks go to my outstanding supervisory team; Prof. Liz Twigg and Dr. Carol Ekinsmyth and the unofficial but very important Prof. Graham Moon, who have provided me with excellent advice and expertise. They have allowed me to realise my plans and I could not wish for a more supportive team to help me.

This thesis could not have been written without the help of ASH Wales. I am hugely grateful for the time and effort the entire ASH Wales team has put into helping me understand the facets of The Filter Wales. Special thanks go out to Emily Cole who showed me how The Filter Wales social media engagements worked and who answered my many questions.

I would like to acknowledge my colleagues at the Geography department at the University of Portsmouth who have made me feel welcome in the UK. I would specifically like to thank my fellow PhDs; Jenny Cox, Lauren Knight, Paul Weber, Margarita Tsakiridou, Regina Kolzenburg, and Taye Famuditi - you were indispensable. Another special mention goes out to Adrian Tear for teaching me SQL and helping me out with coding when I got hopelessly stuck.

Finally, I would like to thank my wonderful family who has been extremely supportive over the three years. I want to thank my mom and dad for their support and letting me live back home while I was writing up the PhD. But, the utmost appreciation goes out to my brother who helped me out with this PhD in all kinds of ways.

I really could not have done it without you.
**Dissemination**


Chapter 1. Introduction

On a global scale, it is estimated that one billion people smoke tobacco and this habit will cause the premature death of approximately half of them (WHO, 2015b). So far, a significant amount of research has been done to understand this damaging behaviour and developed ways to cut down smoking prevalence. This study adds to this vast literature by providing an exploration of what young people say about smoking via social media, namely Twitter. This is important because smoking habits are initiated in youth, and it is at that life stage that interventions to eliminate smoking need to focus. This age group is also most likely to interact and share experiences and opinions on this social media platform which provides a novel approach to researching youth smoking. Ultimately, this study aims to comprehend how better to prevent young people from taking up this deadly habit. This introductory chapter starts with a description of the ‘smoking epidemic’ and the ‘smoking endgame’ on a global level and narrows in on the focus of this study which centres on young people, largely from Wales, a country that is part of the United Kingdom.

1.1 The background

Worldwide, 80% of the one billion smokers are from low- and middle-income countries and smoking prevalence is increasing rapidly in emerging economies in Africa and the East Mediterranean (Bilano et al., 2015). Bilano et al. predict that smoking prevalence will increase to approximately 1.6 billion in 2025 due to the smoking uptake in these countries and a higher smoking prevalence for women in countries across Europe and the Middle East. This prediction is based on a much-used model created by Lopez, Collishaw & Piha (1994) to display the trajectory of the ‘smoking epidemic’ spreading through populations for the future. The original model was created through a temporal analysis of worldwide, national smoking prevalence and was later fitted with a timeline to exhibit the smoking trajectory of high-income countries as shown in Figure 1.1 (Thun, Peto, Boreham, & Lopez, 2012). This four stage model communicates the delay in smoking uptake for low-income countries (compared to high-income nations) and its full effects on mortality (Thun et al., 2012). Nations are thought to transition through these stages. The ‘smoking epidemic’ model shows where cigarette smoking was taken up by men before it caught on for women and depicts a continuum spread over many decades rather than a series of isolated events. The model allows countries to place themselves in one of the four stages to understand how the smoking epidemic manifests itself and what the consequences are. It
reveals how, if smoking prevalence decreases, a rise in smoking-attributed mortality is still likely to continue for some time (Thun et al., 2012).

According to Figure 1.1, the habit of smoking has become less dominant in high-income societies where smoking prevalence has gone down for both men and women since the mid-sixties. The UK is in the fourth stage where the percentage of all smokers is declining but levelling off. Further decreases necessitate a focus on young people starting smoking as it is their input which maintains the prevalence. The low and middle-income countries (mentioned earlier) are in the first two stages, and it is predicted that the smoking prevalence in these countries will rise unless severe anti-smoking action is taken.

Figure 1.1 The ‘smoking epidemic’ model for high-income countries time frames (Source: Thun et al., 2012).

Critiques of the model consist of notions that it does not account for cultural differences, and many countries do not follow the four-stage pattern (Thun et al., 2012). For example, different from the last stage in the model, the smoking-attributed deaths for men and women have converged in countries such as Canada and The Netherlands. Nonetheless, the original model has value in predicting the areas where tobacco control actions are most needed. The World Health Organisation (WHO) uses the model to predict the trajectory of
smoking prevalence and to help communicate the future health effects of current smoking for nations globally (Thun et al., 2012).

1.1.1 The Smoking Endgame

Many countries have acknowledged the smoking epidemic and signed the ‘WHO Framework Convention on Tobacco Consumption’ (FCTC; WHO, 2003). The WHO FCTC is an “evidence-based treaty that reaffirms the right of all people to the highest standard of health” (WHO, 2015a). This underpins the ‘smoking endgame’. The smoking endgame refers to the global effort to reduce the smoking prevalence to 5% or below by an announced date which varies per nation (Eriksen, Mackay, Schluger, Islami Gomeshtapeh, & Drope, 2015; Malone, McDaniel, & Smith, 2014). The realistic target is at least a 30% smoking reduction in adult smoking prevalence by 2025 in each partaking country (Eriksen et al., 2015). The endgame aims to entirely eliminate the health toll of tobacco products by affecting the tobacco industry’s economic interests through anti-tobacco legislation (e.g. prohibit sale and manufacture of tobacco) (Barnett, Moon, Pearce, Thompson, & Twigg, 2016).

By signing the Convention, nations are bound to implement anti-smoking policies and to prevent and control the use of tobacco products (WHO, 2015a). The overall aim of the WHO FCTC is to tackle the smoking epidemic beyond national borders, and the primary objectives can be summarised with the acronym MPOWER;

- Monitor tobacco use and prevention policies
- Protect people from tobacco use
- Offer help to quit tobacco use
- Warn about the dangers of tobacco
- Enforce bans on tobacco advertising, promotion and sponsorship
- Raise taxes on tobacco

A report came out ten years after the start of the WHO FCTC revealing that 179 countries have signed the convention (WHO, 2015a). The review report illustrated further that many countries have seen a decrease in smoking prevalence and mentioned two countries with the most significant decline; 22% in Uruguay and 25% in Turkey (WHO, 2015a). The endgame is in sight, but it will be challenging to reduce smoking prevalence by 30% by 2025 as will be exemplified in the remainder of the thesis.
Nowadays, people are aware of the health risks related to smoking, but that has not always been the case; only since the second half of the 20th century has there been a focus on tobacco control, even though, the connection between smoking and specific health risks had been made decades earlier (J. Hughes, 2003; R. West, 2006). Until the 1970s, no government took serious action against smoking uptake even though the majority of men and many women were already smokers (R. West, 2006). From the seventies onwards, tobacco control policies started with raising awareness of the health risks of smoking and gradually increased to all areas of tobacco control including restrictions on tobacco sales and cessation counselling through local health services (L. F. Stead, Bergson, & Lancaster, 2013). To prevent youth smoking, information-based smoking cessation and prevention programs at schools had started around the same time as the tobacco control policies. Since then, anti-smoking programs have evolved into interventions concentrating on a range of topics including providing information about smoking rates and harms from smoking, teaching youth how to be more socially competent, teaching skills to refuse offered tobacco, and multimodal programmes with parents, teachers, and communities (Grimshaw & Stanton, 2006; R. E. Thomas, Perera, Mclennan, & Perera, 2013).

1.1.2 Smoking in the United Kingdom

This section will zoom in on the UK, where (as in many European countries) smoking is the leading preventable cause of death. In the UK, one out of every five deaths can be assigned to smoking behaviour, and smoking causes the premature death of an estimated 96,000 people a year (ASH, 2016a; Scarborough et al., 2011). Smoking-related causes of death include, but are not limited to, lung cancer, bronchitis, emphysema, and cardiovascular diseases (ASH, 2016a). Premature deaths and preventable ill-health due to smoking produce a substantial economic burden on the healthcare system, and the UK health sector spends approximately 2 billion pounds per year on treatment of smoking-related diseases (ASH, 2015b; Office for National Statistics, 2014b). Research by Action on Smoking and Health (ASH) estimated the costs of smoking for society to be £13.9 billion a year including National Health Service, employer, and environmental costs (ASH, 2015b). Lightwood & Glantz (2016) calculated the reduction of 2.2% in total healthcare expenditure for the USA for 2012 if the smoking prevalence dropped by 10% in every state. Their figures indicate that, with a small change in smoking prevalence, a significant amount of money could be saved even in the short term (Lightwood & Glantz, 2016). Assuming that these conclusions are transferable, anti-smoking initiatives could make a considerable change in the economic burden on the UK society.
For the UK, the numbers from the Office of National Statistics on smoking prevalence in 2015 are 19.9% for men and 18.1% for women. Over the past 40 years, there have been some significant reductions in the prevalence of smokers in the UK as demonstrated in the following figures which reveal the proportion of men and women since 1974 that are smokers (Figure 1.2), who have quit smoking (Figure 1.3), and have never smoked (Figure 1.4). The results illustrate that smoking prevalence for both men and women has halved and the gender gap has almost disappeared (Figure 1.2). The second graph (Figure 1.3), on the number of quitters in the UK, demonstrates how the number of people who have quit smoking increased but a dip is shown after 2010. The graph depicting the proportion of adult non-smokers in the UK (Figure 1.4) reveals that women have always been more likely to be non-smokers, but the gender disparity is decreasing. These numbers indicate that not only fewer people ever smoke, that a significant proportion of the people have quit since the seventies, and that gender differences are decreasing. The UK smoking prevalence profile can be categorised as part of phase four in the ‘smoking epidemic’ model (Bosdriesz, Willemsen, Stronks, & Kunst, 2016; GBD 2015 Tobacco Collaborators, 2017) and these numbers suggest that smoking has become a less common habit in the UK.

Figure 1.2 Proportion of British adults (16+) who smoke (1974-2012) (adapted from General Lifestyle Survey, 2013)
Figure 1.3 Proportion of British adults (16+) who quit smoking (1974-2012) (adapted from General Lifestyle Survey, 2013)

Figure 1.4 Proportion of British adults (16+) who have never smoked (1974-2012) (adapted from General Lifestyle Survey, 2013)

Wales (which is where the research sample used in this thesis originates) does not follow UK trends exactly. The next line chart (Figure 1.5) demonstrates the difference in adult (16 years and older) smoking prevalence between Wales and the UK between 1998 and 2011. Since 1998, Welsh adult smoking has been following the UK average but with a closer fluctuation between men and women. However, from 2008 onwards the UK trends of declining smoking prevalence are continuing, while the Welsh population has increased their proportion of smokers above the UK average. Interestingly, since 2008, the proportion of female smokers is higher than male smokers in Wales which is not in line with the UK trend.
The UK government has implemented programs and a tobacco control action plan to achieve the objectives of the WHO FCTC and reach for a ‘smoke-free future’. However, the number of smokers in 2015 resides around 19% of the UK population and 23% for Wales. Comparing Welsh smokers to their European counterparts reveals that, overall, they score below the European average on smoking prevalence (WHO, 2016). However, Welsh men smoke below the European average, but Welsh women smoke above which indicates a divergence from the rest of Europe (WHO, 2016). Moreover, Wales has a smoking disparity based on social-economic positions (SEP). The Welsh Health Survey (2015) showed that there was a social gradient in adult (16 years and older) smoking in Wales in 2014; smoking prevalence was four times higher for people who never worked or were long-term unemployed (43%) than for people in managerial and professional jobs (11%).

As seen in Figure 1.3, the number of quitters has increased, but it does not specify at what age they individuals quit. Jha (2009) reviewed data on smoking-related mortality and claimed that cessation before middle age (30 to 40 years) avoids more than 90% of smoking-related lung cancer and reduces risks of other related diseases. Ideally, the incentive should be not to start smoking in younger age, but preferably young smokers are induced to quit smoking before the age of 30. Therefore, attention is turned to the younger generation and their smoking uptake, habits and cessation.
1.1.3 Youth smoking in Wales

Approximately 13,000 teenagers start smoking each year in Wales (ASH, 2015c; ASH Wales, 2010). The clustered bar chart (Figure 1.6) illustrates how these smokers are distributed by gender and age in the year 2014. This bar chart is based on two Welsh surveys; Welsh Health Behaviour in School-aged Children (HBSC, 2015) and the Welsh National Health Survey (Statistics For Wales, 2015). A reliable measure of smoking prevalence among teenagers is the HBSC survey, adopted by the WHO Europe as a collaborative study between 42 countries in Europe and North America. It is an annual survey given to children aged 11, 13, and 15. Older young people (16 years and older) are categorised in the national health survey as adults, and their smoking prevalence is taken from the Welsh Health Survey. This bar chart illustrates how in the 11- to 12-year-old age group, less than 1% smokes, and this increases to 3% with teenagers two years older. The 15-year-olds show a gender gap where women smoke more than men, and while this persists into the next age group of 16 to 24, it is reversed from 25 years and up (Figure 1.6). For both genders, there is an apparent increase in smoking prevalence with age, but there is a steeper rise for men.

![Figure 1.6 Smoking prevalence of age of men and women in Wales in 2015](adapted from The Welsh Health Survey, 2015 & HBSC, 2015).

1.2 The problem

As seen in Figure 1.6 above, there is a steep increase in smoking prevalence from age 11 to 24 years-old. Something happens in the transition to adulthood that impacts on whether a person becomes (or remains) a smoker or not. This group of young people is, therefore, a central focus for research. The part smoking plays in their everyday lives is important for
understanding their reasons for smoking and the difficulties they may encounter when attempting to quit.

The WHO FCTC aims to denormalise smoking in society through bans on smoking in public places and informing the public about the health risks. However, Gagne, Frohlich & Abel (2015) discussed the knowledge of smoking health risks for young adults in Switzerland, and their study revealed that their perceptions of smoking were created by life-long socialisation instead. This finding suggests that educating young people about the risks of smoking is less effective than the influence of their family and friends (Gagné, Frohlich, & Abel, 2015). More studies found the connection that coming into contact with smoking during childhood is associated with later smoking (Becklake, Ghezzo, & Ernst, 2005; Fuller et al., 2014; Griesbach, Amos, & Currie, 2003; Hoving, Reubsaet, & de Vries, 2007). To be specific, ASH estimated that each year 23,000 young people in England and Wales start to smoke by the age of 15 due to exposure to smoking in the home (ASH, 2015c).

Another challenge with youth smoking is that young people are considered the key market for tobacco companies to recruit new smokers (Moodie, MacKintosh, Brown, & Hastings, 2008). The tobacco industry lures this group to smoking through appealing (indirect) marketing and attractive smoking images that challenge the ethos of messages from anti-tobacco campaigns and override tobacco control policies. This tactic is clever and effective as adolescence is a time for self-identification and (according to pro-tobacco promotion) looking ‘cool’ and ‘rebellious’ is easily done through smoking (Denscombe, 2001b; Spijkerman, Van Den Eijnden, & Engels, 2007). The more visible and appealing smoking is for young people, the more likely they are to perceive smoking behaviour as normal and socially desirable (Lovato et al., 2010). Therefore, it is the social meaning and the self-perceived image of smoking that is central to an understanding of continued youth smoking.

A key demographic to focus upon are young people between 11 and 25 as 99% of all current smokers have started within this age group (ASH, 2015a). Moreover, this broad youth age group tends to be overlooked as an interesting age range in research. Most research either focuses on children (up to the age of 16) or adults (over 16) as illustrated in the figures above. A better understanding of smoking behaviour in this particular age group is important to understanding how a smoke-free future can be reached, and a successful smoking endgame achieved.
1.3 This study’s justification and aims

It is difficult to find out how young people perceive smoking in their day-to-day existence, how they justify their smoking behaviours, and how they associate smoking with other parts of their lives. By understanding the complex meaning and interpretation young people give to smoking, a start can be made to align the anti-tobacco initiatives to prevent smoking uptake better and more efficiently influence youth smoking cessation.

Essentially, most existing knowledge of engagement and reach of health organisations and our understanding of youth smoking is derived from questionnaires, interviews, focus groups, and surveys. Much less attention is placed on the relatively novel method of utilising social media as a source of smoking data. This is a missed opportunity. First, social media studies have advantages over traditional methods. The challenges of traditional methods are that large datasets provide little depth to the results whereas interviews and focus group provide a more profound understanding but have smaller sample populations. This study resides between the two with a study into the broader understanding through a large number of posts on a social media site (i.e. tweets made via Twitter). Second, many surveys are undertaken in school settings as part of the mandatory curriculum, but social media allows the examination of smoking in a naturalistic setting, popular with young people. This novel method can reveal smoking from the perspective of young people without them being compelled to think about the subject or actively participate in the research. Third, social media is a platform increasingly used by young people to interact and describe their opinions, activities, and updates about their lives (D. J. Hughes, Rowe, Batey, & Lee, 2012). Twitter has a vast array of possibilities that avoid challenges such as socially expected replies and low response rates from young people. Researchers are starting to realise the potential that this information can reveal uncensored aspects of people lives from an unforced perspective and, through these posts, reveal what people find important and their understanding of their lives and actions.

This study cooperates with The Filter Wales, a campaign that engages with young people on the topic of smoking and, with their assistance, this study has the opportunity to explore the utility of social network sites to understand young people’s perceptions on smoking. The Filter Wales campaign (a subdivision of ASH Wales) has an extensive social network engagement to target young people (age 11 to 25) through Twitter, Snapchat, and Facebook. The Filter Wales uses Twitter most extensively by not merely posting information about the possibilities of quitting smoking, but also searching for smoking tweets by young people in Wales and interacting with those young people online. In this
way. The Filter Wales connects with people that may not have been aware of the organisation and helps young people with their smoking activity on a platform they are already using. Young people engaging in the Twitter element of the campaign provides the sample members for this study and their tweets, and Twitter profiles provide de data.

Uniquely, this thesis deliberately studies smoking references on Twitter and young people in contact with The Filter Wales through Twitter data alone to expand the knowledge on youth smoking through a novel platform. Moreover, this platform provides additional information about the reach of the campaign that would be difficult to achieve through traditional channels. What this Twitter-based study explores, is stated in two aims;

1. to understand the reach of the Twitter element of a social media campaign (The Twitter element of The Filter Wales).

2. to assess the text content of tweets about smoking to understand more about the social meaning of smoking and health risk co-behaviours.

This study makes a case for using social media as a significant influence on youth smoking knowledge, and it adds to the literature by expanding on key insights derived from traditional domains of youth smoking research. For the first time, this study combines the efforts of a smoking cessation campaign with the individual insights of young people to better understand youth smoking.

1.4 The structure of the thesis

This thesis is presented in ten chapters (including this Introduction which serves as a foundation to set the scene for the research). Chapter 2 continues with an exploration of the literature base on youth smoking. This is presented according to the different individual smoking stages that a person can transition through, starting with initiation, moving on to social smoking, quitting, and ending with the ex- and non-smokers. Additionally, the second chapter illustrates the place of e-cigarette, marijuana and shisha smoking within this youth smoking narrative and the association of other health behaviours with tobacco smoking in the form of health risk co-behaviours. The focus on different smoking products and health co-behaviours is important as smoking is not done in isolation and, by connecting tobacco use to other health behaviours, it becomes more distinct how people view their own behaviour and treat their body.

The second review chapter, Chapter 3, expands on the initiatives for a smoke-free future in the UK such as the ban of smoking in public places and the regulations on the sales of tobacco and its unintended consequences. These consequences are an essential element in
the realisation that anti-tobacco approaches are not all advantageous to young smokers. The last part of Chapter 3 provides an in-depth account of The Filter Wales, including an evaluation made by Cardiff University and highlights the research gaps in their evaluation that this study will address by researching the Twitter data from the young people in contact with the Filter Wales. Chapter 3 finishes with a detailed description of the aims and objectives of this study.

Chapters 4 and 5 outline the methodological basis for the thesis. Chapter 4 provides a description of the Twitter social media platform and explains the possibilities of Twitter as a research tool compared to traditional methods. The chapter also addresses the ethics of using Twitter in social research. Additionally, Chapter 4 illustrates the gathering and preparation process of the Twitter data and finishes with a guide on how to read the tweets in this study. The second methodology part, Chapter 5, outlines the design decisions for the research and elaborates on the methods of analysis. These methods consist of descriptive quantitative and more in-depth qualitative data analyses and incorporate both traditional and novel approaches for this Twitter-based study.

The next four chapters comprise the substantive results chapters. Chapter 6 begins with an examination of the sample members and the social inequalities that can be derived from the Twitter accounts, including mapping the location of the place of residence of the sample and attaching this information to deprivation levels and a rural/urban classification. Chapter 7 details a quantitative content analysis of smoking-related tweets and places the content of the smoking tweets in various time frames. This chapter also presents a sentiment analysis examining how views on smoking differ with smoking products and through time. Chapter 8 continues with content analysis but takes a qualitative approach to the tweets that include simultaneous engagement in other health risk behaviour (i.e. co-behaviours). Content analysis is applied to uncover how these health behaviours, i.e. alcohol use, healthy eating, and exercise interplay with smoking. The last results chapter, Chapter 9, centres on the examination of a smaller group of sample members and provides an in-depth discussion on how smoking fits the way they use Twitter. This last results chapter also provides perceived marginalisation, more detailed text-derived context on quitting attempts, and contact with The Filter Wales within the Twitter history of fifty sample members.
Chapter 10, the final chapter, concludes the thesis with an overview of the results of this study. The chapter elaborates on the original contributions to knowledge derived from the various analyses presented in Chapters 6 to 9. The academic and policy implications are also discussed, highlighting the importance of the work presented here. The chapter closes with a discussion of the limitations of the study and recommendations for future work.

The following chapter will elaborate in detail on current knowledge about the social meaning of smoking through an in-depth discussion of previous smoking literature on young people.
Chapter 2. Literature review on youth smoking

This chapter is a critical literature review of how young people perceive smoking. It is, therefore, concerned with how smoking is regarded, negotiated, understood, and interpreted and explores how any one individual may hold both positive and negative attitudes to the behaviour simultaneously. For example, young people may regard tobacco smokers as ‘cooler’ than non-smokers, but at the same time, they may interpret tobacco cigarettes as damaging to health. This mismatch, as well as the broad perceptions that an individual holds, may be explained through an understanding of the social meaning of smoking (Frohlich, Poland, Mykhalovskiy, Alexander, & Maule, 2010; Poland et al., 2006; Ritchie, Amos, & Martin, 2010b), referring to how smoking is represented differently in various social settings. Some authors would argue that a full understanding of social meaning is key to the understanding of the diverse sources of resistance to anti-tobacco initiatives (e.g. Lantz et al., 2000; Poland et al., 2006; Proctor, 2015). Accepting that people perceive smoking differently, and exploring how, increases the understanding of youth smoking and the challenges of creating a smoke-free future (De Clercq, Pförtner, Elgar, Hublet, & Maes, 2014).

The overall goal of this chapter is to establish a foundation for the empirical analyses presented in Chapters 7, 8 and 9 where evidence is garnered from the Twitter dataset and used to explore these themes of perceptions across smoking behaviour and other co-behaviours. The review presented here searches to examine the key perceptions linked to social meaning for young people across a framework of stages of smoking behaviour. The framework for this literature review uses the stages of youth smoking adapted from the work of Mayhew et al. (2000). Their stages, which are more or less dynamic, include; non-smoking, smoking initiation, experimenting and regular smoking. In this study, additional stages described as ‘social smoking’ and ‘quitting smoking’ are included to encompass a more comprehensive overview of the possible smoking stages through which young smokers may transition. Particularly, this strategy is valuable to present smoking as a process instead of a static position. Here these stages are described more fully as initiation and experimenting (section 2.1), social (and regular) smoking (section 2.2), and quitting smoking (section 2.3). Non-smokers are discussed towards the end of the review and encompass the group who have never smoked and the group of ex-smokers who have successfully quit smoking (section 2.4).

This review embraces the fact that tobacco smoking is often not undertaken in isolation. Therefore, this chapter also addresses the interplay and interaction of smoking with other
health risk behaviours and again considers them within the framework described above. Cigarette smoking behaviour may interplay with consumption of other smoking products (i.e. marijuana, e-cigarettes and shisha) and other health-related behaviours (i.e. alcohol use, healthy eating, and physical exercise). Patterns of co-behaviour are often interrogated in health surveys to understand the individual characteristics associated with these behaviours. The nuances of their interplay and how social meaning is created and negotiated around them simultaneously are relatively under-researched. Section 2.5 of this chapter, therefore, provides more detail on co-use of tobacco cigarettes alongside e-cigarettes, marijuana, and shisha smoking. Furthermore, the uncovering of these nuances in health risk co-behaviour and their particular attributes that link them to smoking will be covered in the last part of this chapter (section 2.6). Attention first returns to understanding smoking behaviour via the smoking stage models outlined above, beginning with smoking initiation.

2.1 Smoking initiation

This first section of the literature review outlines the first experience with smoking and the elements that influence actual smoking initiation. An important start to the review on smoking initiation are two studies by Kremers et al. in 2004. Kremers et al. (2004) designed a framework outlining motives for smoking in order to reveal the type of non-smokers who are most likely to begin smoking. They categorized young non-smokers according to a timeframe that did or did not include future smoking activity. The types were described as a committer (sure never to start), immotive (no plans to start), progressive (possibly start within six months), and contemplator (start within next six months) (Kremers, De Vries, Mudde, & Candel, 2004). However, in Kremers et al.’s later study that year, the authors applied the theory in a longitudinal self-administrated questionnaire in schools to 10,170 young people across six European Union countries and revisited the initial findings. Here they discovered that smoking initiation is mostly unplanned and young people initiated without having concrete motives (Kremers, Mudde, & De Vries, 2004). These findings indicate that although specific factors influence smoking initiation, young people do not necessarily make an active decision towards that behaviour. Consequently, most of the literature presented in this section is based on retrospective views of smokers about when they started smoking.
2.1.1 Favourable opinions

A favourable opinion relates to the expected image smoking can give to the individual and when an individual perceives that the activity will improve their external image, the odds of initiating smoking are increased (B. N. Smith, Bean, Mitchell, Speizer, & Fries, 2007; Spijkerman et al., 2007). For example, smoking behaviour as depicted by actors in movies leads adolescents to hold more pro-tobacco beliefs and attitudes as “it teaches people that smoking is normal, widespread, and even desirable in society” (Charlesworth, 2005 p.1526). Many other studies have argued that engaging in smoking results in young individuals concluding that they are looking ‘cool’ and ‘like an adult’ and it reinforces positive feelings and opinions regarding smoking (Klein, Sterk, & Elifson, 2013; MacFadyen, Amos, Hastings, & Parkes, 2003; Spijkerman et al., 2007). Spijkerman et al. (2007) showed that smokers thought of nonsmokers as “less well-adjusted, less rebellious, less cool, and less attractive” (p.93) than other smokers. Similarly, Mayhew et al. (2000) claim in their study on adolescent smoking stages that the smoking initiation stage is associated with the overall idea of improving one’s image.

The favourable opinions of smoking portray itself different for men and women. When questioned about magazine pictures of people smoking, young people responded that male smokers were considered tough and rebellious whereas female smokers were depicted as sexy and fashionable (Amos, Gray, Currie, & Elton, 1997). Moreover, there is an often reported perception that smoking helps with losing weight and preventing weight gain and is particularly popular with women to justify their smoking initiation (Alexander, Frohlich, Poland, Haines, & Maule, 2010; Amos & Bostock, 2007; Larsen, Otten, & Engels, 2009). In contrast, for young men who are particularly concerned with body image, the opposite is frequently true. They are less likely to smoke as smoking interferes with lung capacity that is necessary for effective exercising (Amos & Bostock, 2007; Verkooijen, Nielsen, & Kremers, 2008).

These positive connotations with smoking outweigh the perceived risks of ill-health which may be held during this early phase smoking. For young people, there is a sense of being invincible and getting sick from smoking is something that happens to older people (Amos, Wiltshire, Haw, & McNeill, 2006; Heikkinen, Patja, & Jallinoja, 2010; MacFadyen et al., 2003; Peretti-Watel, Legleye, Guignard, & Beck, 2014; Song et al., 2009). Not only is ill-health not relevant during the initiation phase, but it is also possible that young people are ignorant or confused about smoking-related health risks. Song et al. (2009) argued that, for adolescents, a lack of knowledge or concern about the health risks increases the chance of
initiating smoking by three times. Furthermore, Peretti-Watel et al. (2014) showed that among their sample group of French adult smokers, a low perception of health risk, and getting health information from the internet or relatives related significantly to smoking initiation. Overall, the influence that smoking has on identity-formation tend to be positive during this initiation stage, and health risks are ignored, or their relative importance is relegated when young people begin smoking.

2.1.2 Peer influence

An increasing number of studies link friends and peer influence to smoking initiation (De Vries, Engels, Kremers, Wetzels, & Mudde, 2003; Fuller et al., 2014; Mercken, Candel, Willems, & De Vries, 2009; Pierce, Distefan, Kaplan, & Gilpin, 2005). For example, results from the ‘Smoking, drinking and drug use among young people in England’ (SDD) survey, indicate that four out of five (81%) young people in England have a family member or friend who smokes (Fuller et al., 2014). While individuals in the household are an important factor, at the beginning of adolescence, a shift is seen whereby there is a higher influence from the peer environment (Mercken et al., 2009). To sustain friendship circles, non-smokers are pressured to engage in smoking by their peers who are smokers (Denscombe, 2001a). Furthermore, Kuipers et al. (2016) argued in their study on socioeconomic differences in perceived smoking prevalence among adolescents in Europe that smoking initiation is more influenced by the perception of peer smoking rates than the actual peer smoking prevalence. They further discuss that lower educational achievement of the individual and parents who smoke are at the base of higher perceived smoking prevalence and overall likelihood of smoking initiation (Kuipers et al., 2016).

In the study on predictors of smoking initiation by Hoving, Reubsaet & de Vries (2007), young women were particularly sensitive to peer pressure. The authors argued that this was because women perceived that smoking would improve attractiveness to men and would compare positively with other women. They further argued that young men were more likely to regularly smoke if they associated smoking with deviant behaviour and feeling the need to rebel which is also most likely influenced by peers (Hoving et al., 2007).

However, other authors argue that although friendships are an essential factor for youth smoking initiation, this is not just a causal relation (Holliday, Rothwell, & Moore, 2010; Mahabee-Gittens, Xiao, Gordon, & Khoury, 2013; Mercken, Sleddens, De Vries, & Steglich, 2013; Spijkerman et al., 2007). According to Mahabee-Gittens et al. (2013), friendships are vulnerable constructs for teenagers. Parents can avert connections to certain peer groups by preventing their children from hanging out with smokers. A related
argument comes from Spijkerman, Van Den Eijnde & Engels (2007) in a study on smoking friend groups among high-school students in the Netherlands. They argued that children create friend groups based on shared opinions and, therefore, the attitude towards smoking is a point by which they pick their friends. Similar results are found in British studies (Holliday et al., 2010; Mercken et al., 2013) in which the researchers discussed that the initial smoking perceptions influence the type of friends and reinforce the influence of friends on smoking initiation.

2.1.3 Experimenting with smoking

Smoking initiation starts off with an experimental phase and adolescence is a time for experimenting (Amos, Agnus, Bostock, Fidler, & Hastings, 2009; Denscombe, 2001a). In the English SDD survey, young people were asked if they thought it was OK to (one-off) experiment with tobacco use and the majority answered affirmatively (Fuller et al., 2014). Most young people hold the idea that a single cigarette will not make a difference to their health (Pierce et al., 2005).

Experimenting with tobacco can emerge from curiosity which makes young people try smoking, and if it is a pleasant experience or if it evoked a desirable effect, it can turn into actual smoking initiation (McClure, Arheart, Lee, Sly, & Dietz, 2013; Pierce et al., 2005). Current adult smokers mentioned, in discussions about their initial smoking experiences, feeling ‘cool’ and feeling relaxed whilst smoking cigarettes (Bernat, Klein, & Forster, 2012). However, positive experiences are not always necessary. Klein et al. (2013) also studied initial smoking experiences of adult smokers and argued that even though the first cigarette was experienced as negative, nearly three-quarters of smokers had their second cigarette within the month. They concluded that the negative experience led to a continuation of trying cigarettes in order to experience a correct outcome which would be a positive feeling. This persistence would not necessarily have happened if the first experience was considered positive (Klein et al., 2013).

Two articles (DiFranza et al., 2004 and Ríos-Bedoya, Pomerleau, Neuman, & Pomerleau, 2009) measured the relative chance of regular tobacco use among adult smokers by uncovering the initial experiences for positive or negative effects. DiFranza et al. (2004) studied the relative chance of regular smoking based on the first experience. They found that both positive and negative emotional responses increased the chance of regular smoking, but a feeling of relaxation after the first cigarette was associated with regular smoking four times more than feelings of nausea, dizziness and irritation. Similarly, Ríos-Bedoya et al. (2009) found that both pleasant and unpleasant experiences were given as
motives to become a regular smoker but that the association was stronger for the pleasant ones.

All in all, experimenting is an unmissable step in smoking initiation as multiple experiments with friends lead to social smoking and nicotine dependency (Gervais, Loughlin, Meshefedjian, Bancej, & Tremblay, 2006). It is during the experimental stage that young people learn how to handle a cigarette and correctly inhale while they can still make up their mind if smoking is something they want to continue (Mayhew, Flay, & Mott, 2000). Around 22% of the young people between 11 and 15 years old have tried a cigarette at least once and that 3% of the 15 year-olds are regular smokers (ASH, 2015c). These numbers show that about three-quarters of young people do not continue to smoke after they experimented with it.

2.2 Social smoking

Between smoking initiation and becoming regular smokers, people often have a period of infrequent smoking behaviour. This is a transition stage where people can choose the occasion to smoke without the physical need for nicotine (J. S. Rose, Dierker, & Donny, 2010). In literature, this period is commonly referred to in research as ‘social smoking’ (Amos & Bostock, 2007), ‘occasional smoking’ (Heikkinen et al., 2010; MacFadyen et al., 2003), ‘non-daily smoking’ (Berg, Ling, et al., 2012) or ‘intermittent smoking’ (Peretti-Watel, Seror, et al., 2014; Rubinstein, Rait, Sen, & Shiffman, 2014). Besides the different terminology, this behaviour is open to interpretation as, especially when self-reported, it can mean ‘classifying oneself as a social smoker’, ‘someone who’s mainly smoking with others’, or ‘someone who’s only smoking with others’ (Song & Ling, 2011).

The vast majority of young smokers see smoking as a social activity (Amos & Bostock, 2007; MacFadyen et al., 2003) and in this thesis, the irregular smoking behaviour is, therefore, referred to as ‘social smoking’. MacFadyen et al. (2003) illustrated how social smoking (and drinking alcohol) is considered a good way to bond with peers through these activities. Amos & Bostock (2007) add to this by reporting that young people experience smoking as a fun activity that creates a socially and culturally acceptable image. They continue by explaining the gender difference in social smoking; for women, smoking was intertwined into the relationship by deepening the social bond whereas, for men, smoking was a way to feel part of a group (Amos & Bostock, 2007).

The adolescent sample members in the study by Johnson et al. (2003) on tobacco dependence expressed a social dependence in which people smoke to maintain connections
that would fall apart if they did not smoke. Moreover, the researchers discussed smoking as being part of the tradition and fitting with social norms and time-related habits such as having a cigarette at the end of a night out (J. L. Johnson et al., 2003). This desire to smoke goes beyond nicotine dependence and regular use as it is smoking for its social meaning. Moreover, parties and nighttime gatherings have a different set of norms around the social acceptance of smoking compared to norms associated with daytime activities and routines (Nichter, Nichter, Carkoglu, & Lloyd-Richardson, 2010).

Likewise, Rooke et al. (2013) in their study on public smoking after the ban on smoking in public places in Scotland argued that smoking is a key activity in young people’s public (specifically night-) life. The structures of being social and the feeling of relaxation and pleasure enhanced the desire to smoke when they are in social settings (Rooke, Amos, Highet, & Hargreaves, 2013). Here, the perception of health risks is overruled by the social desire to smoke even though some participants in their study reported that they felt bad about smoking afterwards (Rooke et al., 2013). Social smoking forms a sense of inclusiveness to a particular group of smokers, and after the ban on smoking in public places, through necessity, the smoking group becomes physically separated from the non-smokers by having to smoke in designated areas (Bell, McCullough, Salmon, & Bell, 2010). These smoking spaces have both positive and negative consequences as it is a convenient way to meet and bond with other smokers or have a more intimate conversation, but it can also be very uncomfortable or isolating when there is no one else smoking or to bond with in the smoking area (Rooke et al., 2013).

Song & Ling (2011) mentioned that social smoking could be a transitional stage on the pathway from regular smoking to cessation; social smoking is often hard to give up because of the positive aspects of social smoking as outlined above in terms of bonding, relaxation and pleasurable effects. For these reasons, this transition phase may last some time before full smoking cessation occurs (Song & Ling, 2011).

2.2.1 Towards nicotine dependence

Social smoking does not exclude young people from health risks and leads to nicotine dependence (Rubinstein et al., 2014). Research has shown that the process of nicotine dependence can start from the first cigarette and be present before an individual becomes a regular smoker (Gervais et al., 2006). However, young social smokers often believe they are not becoming nicotine dependent if they only smoke occasionally and believe that they could easily go without smoking (i.e. quit) if they so desired (Amos et al., 2006; MacFadyen et al., 2003). Moreover, their perception is that nicotine dependence only
follows if they smoke large quantities of tobacco on a regular basis (Berg et al., 2009; Heikkinen et al., 2010).

According to a study by Rubinstein et al. (2014) on young intermittent and daily smokers, the only difference between the two groups was the number of cigarettes consumed per week. Both types of smokers mentioned similar difficulties of quitting their smoking habit indicating that social tobacco use does not exempt people from nicotine dependency (Rubinstein et al., 2014).

2.2.2 Developing the smoker identity

As mentioned above, there is not much that distinguishes regular smoking from occasional smoking in terms of developing nicotine dependence. There may be slight differences in the total number of tobacco products consumed, but even here the distinction is not that clear, especially when operationalising patterns of consumption amongst young people using official surveys. In researching teens in Health Behaviour in School-aged Children (HBSC), for example, young people would be classified as smokers when they ‘smoke at least once a week’ and this definition, therefore, captures both occasional and regular smoking. In adult smoking research (respondents aged 16 and over), however, more importance is placed on the regularity of smoking in survey definitions. The Welsh Health Survey (Statistics For Wales, 2015), for example, uses ‘daily smoking’ as the activity that defines a smoker.

Clearly, in terms of youth smoking, there are many discrepancies across definitions of regular and occasional smoking. In this thesis, attention is instead directed to the factors that influence young people self-identification as a smoker rather than the type of smoker. Adolescence is a time where young people shape their own identity and presentation of self to others (Denscombe, 2001b). For young people, smoking is a way to construct a social identity (Alexander et al., 2010; Amos & Bostock, 2007) and as mentioned previously, smoking may be perceived as ‘cool’, ‘sexy’, or ‘rebellious’. It is interesting to note, however, that gender differences exist in the ways in which these identities are created and negotiated.

There is high social disapproval of female smoking (H. Graham, 2011; Grogan, Conner, Fry, Gough, & Higgins, 2009). Grogan et al. (2009) discussed in their study on 11-15-year-olds that women felt that their smoking habits were more socially controlled by peers. For example, peers would stress the pressure to look a certain way to be allowed to be part of the friend's group (Grogan et al., 2009). However, female students also felt that smoking
‘threatened’ their physical appeal (e.g. yellow fingers or wrinkles) (Alexander et al., 2010). In contrast, for male students, smoking is seen to accentuate their masculine role and identity which is appealing and does not carry the fear of marginalisation (Alexander et al., 2010). This gender disparity affects the way in which young people negotiate their smoking identity. For men, smoking creates a positive image at most times and occasions, but women need to be more selective about when, where and with whom they smoke in order to avoid marginalisation.

These perceptions concern solely the act of smoking and do not refer to nicotine dependence or the acknowledgement of having an addiction. To avoid this contradiction of wanting to have the identity of a smoker but, not wanting to be seen as nicotine dependent, young people may create a strategy that selectively portrays them as a smoker when it is convenient to do so (Bottorff et al., 2004; J. L. Johnson et al., 2003). This strategy entails only to smoke when it improves the social identity, and therefore smoking tends to occur in the presence of others and induces no need to smoke when they are alone (Bottorff et al., 2004). Young people negotiate the balance between a ‘cool’ smoking-related identity, their desire to smoke at other times, and the desire to not be marginalised.

2.3 The quitting stage

This section covers what motivates people to quit and how that leads to success or failure. First, the smoking literature offers several motives for young people not wanting to quit smoking on a behavioural level. For example, young people rarely perceive themselves as smokers or nicotine dependent and therefore do not need to quit (Berg et al., 2009; McClure et al., 2013; Song & Ling, 2011). Furthermore, people smoke because of its desired effects: a sense of relaxation, relieving boredom, and positive beliefs around weight control (Berg, Ling, et al., 2012; Berg, Sutfin, Mendel, & Ahluwalia, 2012). Giving up smoking would mean having to find other means, where possible, to achieve these perceived positive effects. Besides, young smokers often ‘downplay’ the health importance of their smoking behaviour and use it as a justification to continue and the often used fallacy of noting the ‘many-old-people-who-smoke’ argument typifies this attitude (Heikkinen et al., 2010).

However, the majority of young smokers in the UK do want to quit smoking. According to the ASH Wales youth smoking survey undertaken in 2011 across 11-15 years-olds, 75% of

---

1 A great deal of literature focused on women wanting to quit smoking during pregnancy, but this motive is not further developed here as it relates to external drivers making people quit smoking and not individual perspectives.
young Welsh smokers wanted to quit (ASH Wales, 2011). Among the older age group (16 to 24 years old) in the Health Survey of England 2007, the percentage of men and women wanting to quit smoking was 67% and 76% respectively (Amos et al., 2009). In that same survey, 2% of the 16-19-year-olds and 10% of the 20-24-year-olds reported being ex-smokers. These percentages show an evident gap between wanting to quit smoking and succeeding that will be further expanded on in the next section.

Young people vary in the levels of help that they desire to help them quit. In the previously mentioned ASH Wales survey, two-thirds of young people who reported that they wanted to stop smoking indicated that they would like help with their attempt (ASH Wales, 2011). In the English SDD survey, out of the young smokers who had tried to quit smoking, 50% had tried at least one of the smoking cessation services or nicotine replacement therapy (Fuller et al., 2014). Almost half the smokers in England are interested in using online smoking cessation interventions, and yet only a small proportion of smokers currently use these interventions to support cessation attempts (Brown, Michie, Raupach, & West, 2013).

Vangeli & West (2008) questioned British adult ex-smokers what their motivations for quitting were during their final attempt. Future and current health concerns dominated their responses (28.5% and 18% respectively). Interestingly, Fidler et al. (2011) also questioned adult ex-smokers, and they overwhelmingly reported being afraid of losing the pleasure gained from smoking but they actually felt happier as a quitter than when they were smoking. The final motivation might be quite obvious in retrospect but as the study by Fidler et al. showed taking that step might be a fearful prospect.

Uppal et al. (2013) in their study ‘The forgotten smoker’ discussed why most adult smokers do not succeed at quitting. Their results revealed that 70% of smokers had a high motivation to quit but when this was cross-referenced with favourable attitudes towards smoking and ‘liking to smoke’, the motivation diminished and became something that the smokers felt they ‘ought’ to want (Uppal, Shahab, Britton, & Ratschen, 2013). Here the argument is that the external pressure against their smoking habit was high, but the personal motivation is missing.

Conversely, there might be contextual reasons why people do not want to quit smoking, and anti-smoking initiatives do not tackle the ‘underlying causes’ obstructing changing behaviour (G. Rose, 1985). Smoking is still perceived mainly as a social activity and by quitting the individual risks being the one ‘left out’ because they might lose that connection with their social circle (Amos et al., 2006; MacFadyen et al., 2003). Moreover,
smoking becomes integrated with living in disadvantaged areas and creates a social context in which quitting behaviour does not fit (Poland et al., 2006). This last motive is further explained in section 2.3.2.

2.3.1 Relapsing

Bancej et al. (2007) did a systematic review of smoking cessation attempts among adolescents. The studies they reviewed revealed how there is a high report of relapsing even from non-daily smokers and younger adolescents. Instead, it was commonly found that young people reduced the number of cigarettes consumed after these cessation attempts. This result links back to the observation made earlier that intermittent, social smoking can be seen as a transitional phase between daily smoking and quitting (Song & Ling, 2011).

According to ASH Wales statistics (2010), it takes on average seven attempts to quit smoking. In a Welsh survey, young smokers mentioned a number of reasons for failing including the addictive nature of cigarettes, problems at home, and issues around stress and anxiety (ASH Wales, 2010). A strong cause for relapsing is the physical backlash of nicotine dependency (e.g. feeling depressed) and people who smoke more are more likely to relapse (Fidler et al., 2011; J. R. Hughes, Keely, & Naud, 2004; Kotz & West, 2009). These causes for relapsing are focused on the individual’s responses to quitting tobacco and are important factors, but there are also structural pathways that increase the chance of relapsing.

Using data from 4234 young British smokers who attempted to quit smoking in the previous year, Berg et al. (2010) argue that the exact number of failed quitting attempts is unclear as most quitting attempts go unreported. The study illustrated how 80% of current young smokers have tried quitting smoking at least once but, people just forgot that they attempted to quit when the quitting attempt was short, without proper motive, and most importantly unsuccessful (Berg et al., 2010).

The duration of the quitting attempt in adult smokers is an important factor as was also illustrated by Hughes, Keely & Naud (2004) who did a review on relapse data from studies to demonstrate the ‘relapse curve’. They argued that relapsing is most likely within the first week (40-100%). They conclude that the longer an individual remains abstinent, the more likely the quitting attempt will be successful and after 30 to 50 days, the chances of relapsing have lowered to 10-20% which is the lowest chance found in their study (J. R. Hughes et al., 2004). The duration of the attempt is key in successfully quitting smoking,
but for that, the social circle and context (G. Rose, 1985) need to be supportive of remaining abstinent for longer than a week.

There exists a social inequality based on socioeconomic position (SEP) that influences the quitting success. SEP is a common measure to stratify one’s socioeconomic place in society and commonly refers to educational level, employment type, income or (especially in the case of young people) household affluence. According to Kotz & West (2009) measuring SEP by employment type, the likelihood of adult smokers in England making a quitting attempt is independent of the person’s SEP, but smokers with lower SEP (routine and technical occupations) are half as likely to succeed in their quitting attempt than smokers with high SEP (managerial and professional occupations). They showed besides the social gradient in relapsing that there is an equal use of cessation treatments but that the success of quitting smoking with a smoking cessation aid is unequal and again people with lower SEP are more likely to relapse (Kotz & West, 2009).

The results are based on adults but are expectedly transferable to younger smokers as it indicates more structural reasons for higher relapse prevalence among smokers from disadvantaged positions. Other social determinants influence young people’s smoking and quitting prevalence such as family norms and attitudes (e.g. parents are smokers and early onset of alcohol use) and living in disadvantaged neighbourhoods (e.g. areas with high crime prevalence and poor employment opportunities) (Marmot, 2006; Viner et al., 2012). The family norms and attitudes cause a socialisation process in which smoking is seen as normal and quitting attempts as abnormal behaviour. Moreover, there are specific combinations of smoking with other health behaviours that make a quitting attempt undesirable (this is explained in more detail in section 2.7). The determinant of living in disadvantaged neighbourhoods is further explained in the next section.

2.3.2 The hardcore smokers and quitting

The social context is a strong motive to quit smoking or remain a smoker (Idris et al., 2007). For example, in adult studies, rurality has a small independent effect on persistent smoking and intention to quit (Twigg, Moon, Szatkowski, & Iggulden, 2009). Jarvis et al. (2003) added that the ‘hardcore smokers’ were more likely to see smoking as their main pleasure in life and strongly agreed that they liked it too much to quit (Jarvis, Wardle, Waller, & Owen, 2003).

Von Soest & Pedersen (2014) specifically researched young Norwegian smokers between 2002 and 2010 when new and much stricter tobacco legislation was implemented in the
country. This included prohibiting smoking in bars and restaurants alongside a policy vastly increasing tobacco taxation. Von Soest & Pedersen reported that smoking prevalence declined in this group of young people from 21.7% to 7% and noted that the remaining smokers had some distinctive characteristics such as poor school achievement, alcohol and drug use, and living in more disadvantaged areas. The authors claimed that this group had become part of the ‘hardcore smokers’ and would probably remain smokers regardless of tobacco control efforts (von Soest & Pedersen, 2014). These hard-to-engage smokers (regarding cessation attempts) have no intention or real desire to quit which leads to reduced rates of successful quitting and low impact of the offered interventions.

There is very little research on the particular challenges surrounding quitting for young people who smoke and reside in rural areas. Hutcheson et al. (2008) have noted the particular challenges faced by adults, and these challenges are probably similar for young people. The challenges include a lack of local cessation programs, lack of physician visits, financial difficulties, and less control on smoking restrictions. Moreover, the people in rural areas are less aware of the growing body of hotlines, online resources, and national programs to help people quit. This lack of knowledge caused smoking cessation aids not to reach this group (Hutcheson et al., 2008). In urban areas, lack of support services does not seem to be an issue. However, social and cultural norms are imperative for quitting attempts, and the proximity to other smokers hinders a successful outcome (Idris et al., 2007). The individuals’ desire to live a healthier existence is overpowered by structural forces that keep them smoking.

2.4 Non-smokers

The group that has been overlooked so far are the people that do not smoke. This is currently about 80% of the adult population in the UK (Office for National Statistics, 2011). In the English SDD questionnaire, 82% of 15-year-olds had never tried a cigarette (Fuller et al., 2014) and a youth smoking survey by ASH in 2015 illustrated how out of the 11 to 15-year-olds 75% of the young people had never tried smoking (ASH, 2015c). This group forms a key feature in understanding why fewer young people start smoking and how to increase that number.

Mayhew et al. (2000) argued that young people who have never smoked have most likely never thought of smoking as a pleasurable activity and have, therefore, never contemplated starting. Furthermore, young non-smokers are not exposed to positive incentives to start or are capable of resisting and ignoring pressures to smoke (Mayhew et al., 2000).
From the charts (Figures 1.3 and 1.4) in the previous chapter, there has evidently been a stable increase in non-smokers in the population as both the number of never smokers and quitters have increased by at least 15% between 1974 and 2012. Interesting to note is that the boundary between these two groups of non-smokers (never smoked and quitters) may be fluid. Tombor et al. (2015) researched how long it takes for quitters to name themselves non-smokers and found that people who quit smoking do not consider themselves ex-smokers for long. Dependent on their smoking habit (i.e. how much and how long), 80% of the younger ex-smokers continue to see themselves as non-smokers within a year (Tombor, Shahab, Brown, Notley, & West, 2015).

2.5 Tobacco co-behaviours

Attention now turns to the issue of co-behaviour between tobacco consumption and other smoking products. Understanding the social meaning and everyday practice of tobacco consumption often involves interplay and interaction with other consumption patterns and behaviours of e-cigarettes, marijuana, and shisha.

2.5.1 Tobacco and e-cigarettes

Electronic cigarettes have only recently (since 2004) been available on the market and are advertised as a quitting smoking product. They have some similar features to tobacco cigarettes (such as the shape) but, consisting of propylene glycol, vegetable glycerin, flavouring, and (if desired) nicotine (U.S. Department of Health and Human Services, 2016). E-cigarette use is often promoted by manufacturers and marketers as a healthier alternative to cigarettes (Noland et al., 2016). However, the study by Noland et al. also discussed that the switch from tobacco to e-cigarettes is not always to quit smoking but also to progress to a less marginalised form of nicotine intake.

For teenagers, this novel product is more popular due to their availability (as regulations are less strict than for regular tobacco). Moreover, e-cigarettes taste milder, and there are options to add flavour (Goldstone, Macey, & Cass, 2016; G. Moore et al., 2015). Interestingly, recent evidence from young people in Wales reveals that e-cigarette use has increased to be the first experience with smoking, and this experimentation with e-cigarettes has become more popular than experimenting with tobacco cigarettes (De Lacy, Fletcher, Hewitt, Murphy, & Moore, 2017; Goldstone et al., 2016). When surveyed about their reasons for experimenting with e-cigarettes, the main explanations young people in Wales provide are; curiosity (48.7%); because friends are smoking e-cigs (40.1%); and out of boredom (30.7%) (Goldstone et al., 2016). De Lacy et al. (2017) studied e-cigarette use
by Welsh 11-16-years-olds and discovered that there was no evidence of co-use. In contrast, a systematic review by Wang et al. (2016) illustrated that tobacco smokers were more likely to try e-cigarettes and this effect of co-use was higher in adolescents than in adults.

E-cigarettes are a recent invention, and the studies focusing on pathways between e-cigarettes and tobacco do not surpass a few years. Thus far, these studies on e-cigarettes leading to tobacco smoking among young people in the UK have shown no indication of this pathway (Bauld et al., 2017; De Lacy et al., 2017). However, a meta-analysis of nine longitudinal studies in the USA has shown an increased risk of tobacco smoking onset if the young person smoked e-cigarettes regularly (Soneji et al., 2017). Soneji et al. (2017) discussed how the e-cigarette mimics the behavioural scripts of regular cigarette smoking and becomes, therefore, an appealing start to tobacco.

As e-cigarette smoking is seen as less harmful and there are fewer regulations than for tobacco smoking, there is a concern that it might lead to a renormalisation of smoking in public (Mckeganey, Barnard, & Russell, 2016). When young people are surveyed on this topic, their answers suggest that as long as e-cigarettes look slightly different and are not associated with an attractive image, normalisation would not occur according to non-smoking participants (Mckeganey et al., 2016). Wadsworth et al. (2016) researched the use of e-cigarettes and how their use differed from tobacco smoking. Contrary to tobacco smoking, people experienced a lack of capability where they were unsure about the health risks of e-cigarettes. Moreover, e-cigarette smokers felt uncertain about the social acceptability as they could smoke it inside at more places but still induced social disapproval (Wadsworth, Neale, McNeill, & Hitchman, 2016). The authors further discussed how a motive for e-cigarettes smoking was to help in quitting attempts and that the resemblance of the e-cigarette to a tobacco cigarette helped initial attempts at quitting but later the reversed effect was present as the action of e-cigarette smoking reminded the individuals too much of normal tobacco smoking (Wadsworth et al., 2016).

**2.5.2 Tobacco and marijuana**

Tobacco and marijuana co-use show different interactions and interplay than e-cigarettes as marijuana is often consumed with the addition of tobacco and in social settings either in a place of their own or outside (Dunlap, Johnson, Benoit, & Sifaneck, 2005). In section 2.2 was argued that tobacco smoking is a social activity that helps individuals to ‘fit in’ with their peer group and it is this social aspect of smoking which is often seen in marijuana co-use.
Marijuana contains the dried and shredded leaves and flowers of the cannabis plant which can be smoked or ingested. Marijuana is often combined with tobacco in a cigarette (i.e. ‘joint’ or ‘spliff’). A ‘joint’ using 100% marijuana leaf, for example, may be considered too expensive or too intoxicating to smoke without mixing it with tobacco (Bélanger, Akre, Kuntsche, Gmel, & Suris, 2011). Likewise, marijuana resin is often mixed in a joint with hand-rolled tobacco (Haines-Saah, Moffat, Jenkins, & Johnson, 2014). There are several reasons for combining tobacco with cannabis leaves and include making the experience last longer, keeping costs down, and lessening the extreme effect of pure cannabis (Akre, Michaud, Berchtold, & Suris, 2009; Amos, Wiltshire, Bostock, Haw, & McNeill, 2004). With tobacco regulations for minors, it seems to have become easier for them to get their hands on marijuana than tobacco (Tyler, 2015).

The qualitative studies of Akre et al. (2009) and Amos et al. (2004), on tobacco and marijuana co-consumption amongst young people, showed how while many participants smoked marijuana with tobacco, not all of them smoked tobacco cigarettes independently of such joints. This has been shown to influence smoking identity and for those individuals who consume tobacco via mixing it with cannabis may not necessarily self-identify as a tobacco smoker (Akre et al., 2009; Amos et al., 2004; Tyler, 2015). Tyler’s study (2015) on co-use also indicated that young co-users perceived the restrictions on marijuana but, they did not need to travel far to a location where they could buy and smoke marijuana. Moreover, young co-users take the legality of tobacco to their advantage and smoke marijuana out in the open (Tyler, 2015). These findings illustrate that the social meaning of tobacco smoking is defined by its normality in everyday life and it can easily be extended to the co-use of marijuana.

2.5.3 Tobacco and shisha smoking

An inherently social form of smoking is shisha smoking. Shisha is the tobacco product that is used in a waterpipe or hookah for smoking through water damp (Martinasek, McDermott, & Martini, 2011). Shisha is smoked in specific bars or at parties, and the instrument is passed between people for a shared smoking experience. Shisha smoking has been around for many years and was once mainly seen in the Middle East, parts of Asia and Africa but has increased in availability and use in the UK (Akl et al., 2015).

In general, people believe shisha smoking as a more entertaining and a less harmful alternative to cigarettes (Jawad, McIver, & Iqbal, 2014). Moreover, the reasons young people give for smoking shisha are that they perceive it to be less addictive and they report that it looks ‘cooler’ than cigarettes (Akl et al., 2013). However, research on the smoke
quality of both cigarettes and shisha show that the smoke from cigarettes and waterpipes are equally harmful, but the regulations of shisha smoking are not on the same level (Cobb, Shihadeh, Weaver, & Eissenberg, 2011).

The younger generation in the UK is more likely to have experimented with shisha than with cigarette smoking (Akl et al., 2015). Among young English respondents from the SDD survey, 10% of the respondents have smoked shisha at least once, and out of the cigarette smokers, 56% have tried shisha. Less than 1% of all the questioned young people said they were regular shisha smokers (Fuller et al., 2014). Shisha smoking is a social activity and not commonly used on a daily basis. Smoking tobacco through a different medium such as a waterpipe can still result in nicotine dependence. There is evidence to suggest that occasional shisha smoking may drive people towards cigarette smoking to satisfy a nicotine craving when not in reach of a waterpipe (Fielder, Carey, & Carey, 2013; Martinasek et al., 2011).

2.6 Tobacco and other health-related behaviours

The above sections highlight uncertainties concerning the interplay between smoking and other smoking products across a number of themes including pathways, quitting attempts, (assumed) healthy alternatives to tobacco smoking and social meaning. It is important to recognise these complex interactions to understand current youth smoking behaviour more fully. Attention now moves to discussing the interplay of tobacco smoking and other health behaviours in the following section.

When considering health damaging or health-promoting behaviours, they are very rarely undertaken in isolation. Tobacco consumption is no exception and smoking is often undertaken in social or other settings alongside practices that may also damage health but in some instances, alongside practices that promote health and sustain wellbeing. Young people are often aware of how such activities may influence health in either a positive or negative way (van Lenthe et al., 2009), but as with the case of tobacco consumption, the rational choice to undertake such activities is influenced by social practices and norms of behaviour (i.e. social meaning). It is not the intention here to describe social meaning for all possible other health-related behaviours. Instead, the section highlights how other behaviours interact with smoking to influence young people’s perceptions of the social meaning of tobacco consumption. The other health behaviours that are given specific attention in this section are alcohol use, healthy eating and physical exercise together with smoking are defined as the ‘holy four’ (Martin & McQueen, 1989). They often take centre stage in surveys on health-related behaviours because of their direct effects and long-term
influences on health and wellbeing (see HBSC, 2015; Lakshman et al., 2011; Statistics for Wales, 2015).

Two systematic reviews on health co-behaviours by Wiefferink et al. (2006) and Peters et al. (2009) have shown how the uptake (or not) of smoking, alcohol use, healthy eating, and physical exercise across individuals manifest as a set of regular patterns; smoking and alcohol consumption have a negative correlation with healthy eating and physical activity in that the people who eat healthy foods and undertake regular exercise are less likely to smoke or consume an abundance of alcohol. The authors also considered reasons why young people engage in unhealthy behaviours and they found that while they are aware of health risks and the benefits that abstinence may afford, they acknowledged the immediate short-term gains from engaging in the unhealthy behaviours such as having a cigarette or alcoholic drink in order to relax (Peters et al., 2009; Wiefferink et al., 2006).

The Welsh government distributed the health-related lifestyles survey to 12,000 adults (16 years and older) in Wales uncovering the figures on self-reported smoking, excessive alcohol intake, healthy eating, and physical activity. The results of the survey in 2014 revealed that excessive drinking is very prevalent among Welsh adults, that physical activity and healthy eating is slowly increasing, and that self-reported smokers are becoming less frequent. The results further illustrated that smoking, unhealthy eating and physical inactivity reveal a gradient with the lowest prevalence in the most affluent areas and this increases with each level of deprivation. In contrast, the prevalence of drinking alcohol above guidelines is highest in the least deprived area and decreases which each quintile of deprivation (Statistics For Wales, 2015). The results of the survey incorporate young people (aged 16 and above) and is most likely a decent presentation of the health behaviour distribution in Wales.

Very rarely do surveys question about the circumstances of co-behaviour or explain why these co-behaviours emerge together. Moreover, health behaviours play an important part as a protective or indulging element to smoking which differs with the individual smoking stage. The next section illustrates the ways in which co-behaviour may influence smoking in each of the stages of smoking as outlined above. By doing so, it will become clear that the health co-behaviours also play a part in the social meaning of smoking.
2.6.1 Health co-behaviours and smoking initiation

Research shows that at the onset of smoking, alcohol may have a promoting effect. Across all age groups and both genders, alcohol consumption has often taken place during a period of smoking initiation (Chen et al., 2002; Jackson, Sher, Cooper, & Wood, 2002; C. C. Johnson, Webber, Myers, Boris, & Berenson, 2009). In the English SDD survey, 90% of the 15-year-olds who said that they smoked in the last week also drank alcohol and the people that said they drank alcohol in the last week were 19 times more likely to be smokers (Fuller et al., 2014). In effect, most smoking initiation would not have happened without alcohol consumption because both of them reduce inhibitions to experiment with the other (Little, 2000). Furthermore, regular (binge) drinking increases the chance of smoking initiation and regular smoking among young people (Jackson et al., 2002; Leatherdale & Ahmed, 2010; O’Loughlin, Karp, Koulis, Paradis, & Difranza, 2009; Pérez, Ariza, Sánchez-Martínez, & Nebot, 2010).

In contrast, healthy eating and physical exercise are often seen to have a protective quality over smoking initiation. Young people that eat healthily and exercise are less likely to start smoking (Mistry, McCarthy, Yancey, Lu, & Patel, 2009; Wiefferink et al., 2006). Audrain-McGovern, Rodriguez & Moss (2003) discussed how physical exercise exhibits the same properties as smoking by relieving stress and extensive physical exercise requires clean lungs. Therefore, smoking seems less appealing to young people already exercising (Audrain-McGovern, Rodriguez, & Moss, 2003). However, in another study by Rodriguez & Audrain-McGovern (2005), a negative body image was associated with dieting, physical exercise and smoking initiation because all three are considered ways to lose weight. Young people may start smoking to reduce their appetite as well as reduce calorie intake and undertake exercise. They discussed how the focus of poor body image should be more actively connected to physical exercise, rather than smoking, to address issues surrounding poor self-image and smoking initiation (Rodriguez & Audrain-McGovern, 2005).

2.6.2 Health co-behaviours in social settings

Not only is alcohol often implicated in smoking initiation as outlined above but the interplay of smoking and alcohol use are an integral part of their socialising process (C. C. Johnson et al., 2009) and both behaviours are often undertaken in social settings such as pubs, clubs or at home to intensify the bond between peers (Huang et al., 2014). Therefore, it is not surprising that co-users of alcohol and tobacco tend to have co-using friends (C. C. Johnson et al., 2009; Leatherdale & Ahmed, 2010). Moreover, alcohol and smoking are elements that play a key role in the shaping of nightlife activities (Nichter et al., 2010;
Rooke et al., 2013). Nicter et al. (2010), for example, questioned college students about the reasons for their co-use of smoking and alcohol. Their results generated five motives; (1) co-use made social interaction easier, (2) it ‘fits’ with the party scene, (3) there was a perception that alcohol led to less negative side-effects of the tobacco, (4) smoking tends to calm people down when they are drunk, and (5) keeps people awake for the duration of the party. All these motives increased the social experience and the desire to replicate the co-behaviour.

In health behaviour literature, healthy eating and smoking are opposite behaviours and do not combine in social settings in the way that alcohol does. There is something to say for people eating unhealthily in social situations such as parties, but then alcohol consumption is the intermediate (Mistry et al., 2009). One study did find a connection between physical exercise and smoking. Heikkinen et al. (2010) observed in their study of Finnish adult smokers that their participants felt that the health risks of smoking could be balanced out by physical activity. The co-behaviour seemed a way of doing both without undergoing the consequences of smoking (Heikkinen et al., 2010). In this sense, people only smoked in social settings and exercised afterwards too to render the consequences and to stay as healthy as when they were not smoking.

2.6.3 Quitting smoking and health co-behaviours

Continuing on to the quitting smoking stage, people are far less likely to want to quit or reduce alcohol intake than smoking (Statistics For Wales, 2016). Moreover, alcohol is a familiar companion of tobacco and people find it difficult to quit one but not the other (Beard et al., 2016). Alcohol hinders the quitting attempt, leading to a number of adults in Wales stating in the lifestyle survey that they have quit smoking but relapse when they are drinking alcohol (Parry, Carnwell, Moore, & Murphy, 2010).

Dieting and smoking are simultaneously utilised by women who want to lose weight (Larsen et al., 2009; Peters et al., 2009). This makes quitting smoking more complicated as these smokers need to find other activities to remain the same size and replace the emotional regulating effect of smoking. Women mentioned a way to quit smoking is to exercise more so that the body gets healthy and the mind distracted (Amos & Bostock, 2007). Furthermore, smokers over the age of 12 who exercised were more likely to successfully quit (DeRuiter, Faulkner, Cairney, & Veldhuizen, 2008). That success was related to the attribute of physical exercise to ‘clear your head’ and being distracted from the desire to smoke. This study implies that especially for young people, physical activity
should be incorporated in the anti-smoking interventions to prevent experimentation, escalation, and failed cessation attempts (DeRuiter et al., 2008).

This section has shown that the interplay of health risk co-behaviours is not a straightforward association but a complex intertwining of behaviours that are associated with specific activities and requirements.

2.7 Concluding remarks

This chapter has reviewed the literature on the meaning young people give to smoking to better understand the perceptions at each individual smoking stage. The reviewed approaches tend to be extensive surveys of the prevalence of smoking with occasional insights into the concurrent prevalence of behaviour (which leaves very little literature on pathways) or qualitative studies which provided more insight into processes (but had small samples and focused on particular settings and locations).

Most of these studies were snapshots into the lives of young people and their experience with smoking at the moment of the study. This cross-sectional approach prohibits a view of the smoking process as it moves along where tweets from the same individuals can show how the young people feel differently about smoking as they get older. Moreover, the literature here suffered from bias because young people are aware that they are being researched and may answer in a false manner (e.g. socially acceptable answers). Moreover, the literature in this review was based on data that heavily relied on the retrospective accuracy of (young) people, especially for the smoking initiation and quitting smoking section. This recall bias is a challenge for presenting a picture of the young people’s perceptions of smoking. Therefore, a different type of data collection could help with illustrating the motives and meaning as it happens.

This Twitter-based study provides an extensive way of gathering snippets of qualitative information and uses the information given at the time of the event, e.g. people tweeting they had a cigarette and a drink on a night out or tweeting that they quit smoking because of the money problems. Furthermore, in terms of co-behaviour, the specifics to the linkages have been under-researched. These studies have given some reasons for co-use including; marijuana is too strong when it is consumed pure, and e-cigarette and tobacco co-use are a transition from smoking to quitting. Moreover, health co-behaviours have different associations including combining dieting and smoking to lose weight and drinking alcohol and smoking to increase the social experience. In most cases, these studies have not studied the interplay of this co-behaviour. It is difficult to uncover exactly
how these behaviours are combined, but through the Twitter data, linkages can be made visible easier.

This chapter sought to identify the context for youth smoking to understand the perceptions of young people on these topics. This was purely from the young people’s point of view and their individual journey through the smoking stages which brought insight into the unplanned ways in which young people initiate and participate in smoking behaviour under the influence of external forces. It did not go into the anti-tobacco movement or the impact of these approaches on youth smoking. The following chapter will elucidate on the current approaches, policies and interventions to stop people from smoking, de-normalising smoking in public life, and the effects of these interventions on the prevalence of youth smoking. Moreover, the chapter provides an account of the youth-dedicated smoking cessation organisation The Filter Wales, their efforts to help the hard-to-engage young smokers across Wales, and their connection to the present study. The next chapter concludes with the aims and objectives.
Chapter 3. Youth smoking regulations and interventions

This chapter reviews the tobacco control policies, anti-tobacco campaigns and interventions designed to help young people quit smoking or persuade them not to start. The objective of the chapter is to assess the levels of engagement and reach of tobacco control measures among young people. Levels of engagement and reach relate to the likelihood of young people conforming to the tobacco control actions such as only smoking in designated smoking areas and signing up for quitting smoking services. Specific attention is given in this chapter to The Filter Wales; the organisation that provides the core research data for this thesis.

The chapter comprises five sections. The first section (3.1) examines tobacco control policies designed to prevent young people from smoking and covers the control of tobacco sales, legislation to create smoke-free places, and anti-tobacco advertising campaigns. This is followed by a review of health promotion campaigns and programs targeting youth smoking (section 3.2). Having covered this background, section 3.3 develops a critique of existing approaches to the regulation of youth smoking, and the next section (3.4) considers the approach pursued by The Filter Wales. The chapter concludes in section 3.5 with a presentation of the research aims and objectives for the thesis. These aims and objectives are drawn both from this present chapter and from the material presented in Chapter 2.

Anti-tobacco initiatives first began in the 1970s and have taken various forms over time from restricting tobacco sales and advertising to the ban on smoking in enclosed public places that occurred in England and Wales in 2007 (McNeill, Raw, Whybrow, & Bailey, 2005). Since the enactment of the WHO Framework Convention for Tobacco Control (FCTC) (WHO, 2015a), these measures are increasingly focused on the possibility of a smoking ‘endgame’. Although the exact definition of such an endgame varies from nation to nation, the UK has targeted an adult smoking prevalence of less than 5% (Malone et al., 2014). Both tobacco control policies and endgame notions are generic and population-wide. Still, much of their underpinning discourse focuses on smoking prevention amongst young people (Eriksen et al., 2015). The following sections draw out this dual focus.
3.1 Tobacco control policies

Tobacco control policies for young people are designed to make smoking unattractive, complicated and costly by, for example, regulating the sale and price of tobacco as well as restricting the public areas where people can smoke. This, in turn, helps to denormalise smoking by removing the practice from public view.

3.1.1 Reducing tobacco sales

The UK government began to implement regulations on the advertisement of tobacco products in the early 1970s. However, the implementation of these regulations was not monitored, and there was no strong proactive approach towards their implementation (McNeill et al., 2005). A shift in tactics occurred during the 1990s when more emphasis began to be placed on a population health promotion approach to reduce smoking prevalence. This shift was linked to concerns about socio-economic inequalities (e.g. Eriksen et al., 2015; McNeill et al., 2005; WHO, 2015). At the same time, tobacco control legislation and policies multiplied. Illegal sales law was strengthened (1991), the Government committed to a 5% above inflation tax increase for tobacco products (1997), evidence-based guidelines on smoking cessation were established (1998), and the first English NHS smoking cessation treatment service commenced (2000) (McNeill et al., 2005). Regulations designed to lessen the harmful effects of smoking were also gradually placed on tobacco products. For example, in 2001, regulations were passed on the maximum upper limit of tar, nicotine and carbon monoxide for cigarettes sold in the European Union (ASH, 2016b). The main feature in the following decade was the creation of the WHO FCTC in 2003 (see Chapter 1 section 1.2 for the specific goals). Nations that have signed the treaty have had to commit to the implementation of specific tobacco control measures, and such measures include preventing young people from smoking.

One of the main legislative changes came in 2007 by making it illegal to sell tobacco products to anyone under the age of 18 (it was 16 previously). Fidler & West (2010) illustrated how this change decreased smoking prevalence among 16 and 17-year-olds from 23.7% to 16.6%, revealing that even small legislative changes impact smoking behaviour. This point-of-sale legislation was supplemented in 2011 by a ban on selling tobacco from vending machines. The protection of young people against smoke damage intensified even further when, in February 2014, legislation was amended to make it an offence for an adult to buy tobacco or e-cigarettes for minors (known as by-proxy purchasing) (ASH, 2016b).
Since 2008, there have been picture warnings and texts on tobacco products, and legislation relating to product display has been extended via the introduction of plain packaging throughout the UK from May 2017. The WHO theorises that by using plain packaging smokers stand less favourable towards their own tobacco brand and the pictorial health warning may be more effective (WHO, 2011). The packaging is particularly important for youth-focused tobacco control because it removes the possibility that cigarettes can trade on brand images and impulse buys (Uppal et al., 2013).

All tobacco sold in the UK also needs to have the ‘UK Duty Paid’ marking to distinguish it from illegal tobacco (ASH, 2016b). Illegal tobacco sales are a major concern for youth smoking as the sales of these cigarettes are unregulated, and the content of these cigarettes is unknown and possibly dangerous (M. Stead et al., 2013). Most crucially, however, over the last few decades, the cost of cigarettes has increased enormously. This is seen as a critical deterrent for young people. At present, the taxation on tobacco products counts for about 77% of the price of a packet of premium-priced cigarettes (ASH, 2015b).

There is, however, debate about this deterrence. Instead of lowering smoking prevalence, some argue that the increased tax on tobacco products has created a change in smoking behaviour whereby more people smoke economy brands, Roll-Your-Own (RYO) and illegal tobacco to avoid high costs (Rothwell, Britton, & Bogdanovica, 2015). Since the increased taxation on regular tobacco cigarettes, economy brand and RYO cigarettes have been widespread among smokers in Wales and has become even more widely used as a cheap form of cigarette smoking (Gilmore, Tavakoly, Hiscock, & Taylor, 2015). Gilmore et al. (2015) found that among 16-to-24-year-old smokers, 76% smoked cheap cigarettes (48% economy brands and 28% RYO).

3.1.2 Smoke-free places

One of the key recommendations of the WHO FCTC is the need for policies that prevent people from smoking in public and thereby causing harm to non-smoking others. In March 2007, smoke-free legislation came into force in the UK (and many other EU-countries) to protect people from exposure to tobacco in indoor public places and appropriate other places (Levy, Currie, & Clancy, 2013; G. Moore, Currie, Gilmore, Holliday, & L. Moore, 2012). There was an addition to the smoking in public places ban in October 2015 that prohibited smoking in private vehicles carrying children. Currently, efforts are being made to further reduce smoking in areas frequented by young people such as beaches, sports fields and playgrounds but for now, these are based on voluntary initiatives (ASH Wales, 2016; WHO, 2015a).
One of the most common places for young people to smoke is at school with their friends (Cole, Leatherdale, & Burkhalter, 2013). It is, therefore, believed that anti-smoking regulations at schools should have a high influence on the prevalence of youth smoking in that environment. As literature has shown, enforcement of a no-smoking policy at school seems to have remarkable effects as the smoking rates of their students dropped and resistance of their non-smoking students towards smoking increased (Kuipers, de Korte, et al., 2015; Lipperman-Kreda, Paschall, & Grube, 2009; L. Moore, Roberts, & Tudor-Smith, 2001; Pinilla, González, Barber, & Santana, 2002). In a study on schools in Wales in 2001, the researchers concluded that, especially for young daily smokers, a robust anti-smoking policy at schools significantly decreases the number of smoked cigarettes even after adjusting for gender, age, and parents and friends smoking (L. Moore et al., 2001). The policy is especially influential if the school had a zero-smoking policy (meaning young people could not smoke anywhere in the facility). The study by Kuipers et al. (2015) on zero-tolerance school tobacco policy in Europe paid particular attention to the possible differences that educational level may have on the success of such active anti-smoking policies and found that all students, irrespective of educational attainment, would benefit equally from strong policies. Moreover, a no-smoking policy on colleges campuses in California revealed that, when smokers had to smoke outside of the college premises, this drastically decreased the intention-to-smoke among students and made smokers smoke less (Fallin & Roditis, 2015).

### 3.1.3 Prohibiting tobacco advertisement

The Tobacco Advertising and Promotion Act 2002 prohibits the advertising and promotion of tobacco products including tobacco company sponsorship in England, Scotland, Wales and Northern Ireland. From 2012, displaying tobacco products or showing the price of tobacco in shops has been prohibited (Welsh Assembly Government, 2012). The banning of pro-tobacco advertisement has been one of the critical points in the WHO FCTC, but not all countries have signed the agreement, and even when they have done so, resources are not provided to ensure the bans take place. In many countries with high smoking prevalence, such as Indonesia, advertisements targeting young people are still being used. These advertisements are designed to convince young people to smoke as they are the new generation of consumers for the tobacco companies (Eriksen et al., 2015). An example below is a current advertisement from Indonesia evidently aiming at a youth market (see Figure 3.1).
In the UK, according to interviews with young people by Parry et al. (2010), young people noticed a difference in their social life with no smoking in bars and less tobacco advertisement and were less confronted with smoking in their daily lives (Parry et al., 2010). Moodie et al. (2008) examined the perceived prevalence of smoking and the relationship with tobacco awareness in British young people by performing three questionnaire surveys around the time of change concerning point-of-sale displays (one before, one at the time of the point-of-sale ban, and one after the ban). They concluded that awareness of tobacco marketing decreased after the prohibition, which makes the anti-smoking policies to some degree effective (Moodie et al., 2008). Another study on the English tobacco point-of-sale display ban revealed that respondents were less aware of tobacco being sold and the uptake of smoking decreased (Bogdanovica, Szatkowski, Mcneill, Spanopoulos, & Britton, 2015). Similar results came from a study in Ireland where, especially for young people, the point-of-sale ban reduced the awareness of tobacco by a quarter, which indicated a de-normalisation of tobacco sales (Mcneill et al., 2010).

Besides banning pro-tobacco advertising, governments are encouraged by the WHO to display anti-tobacco adverts. Sims et al. (2014) measured the impact of these types of advertisements (in both commercials and posters) in England between 2002 and 2010 and concluded that when people are exposed to anti-tobacco advertisement at least four times a month, this leads to a reduction in consumption. Another study focused on how long the exposure to an anti-smoking commercial lasts among their study population of 18 to 24-year-olds and the authors calculated that the impressionable effect lasts seven days. After this period, intention to smoke immediately returned to the original level (Setodji, Martino,
Scharf, & Shadel, 2014). Long-term exposure to anti-tobacco messaging is most effective but time and money constraints make it difficult to sustain (La Torre, Chiaradia, & Ricciardi, 2005). This anti-tobacco messaging is also challenged by the tobacco companies. They have disparaged advertisements by claiming that the health messages have been overemphasised and the health penalty is not as bad as it is portrayed (E. A. Smith, 2007).

3.1.4 Indirect marketing as a response

Increasingly, to reach out to young people, tobacco companies advertise their products through indirect marketing (see Figure 3.2 below for an example) (Gilmore et al., 2015). Although the advertisement ban was amended in 2008 to include adverts via the internet, new (social) media are a fundamental vehicle; the tobacco industry relies on the youth smoking market for its long-term survival and uses the popularity of (online) entertainment amongst this demographic to its advantage (Moodie et al., 2008). Even with an internet advertisement ban, young people are exposed to pro-tobacco imagery through entertainment channels such as YouTube and mainstream TV series and movies (Cranwell et al., 2014; Leonardi-Bee, Nderi, & Britton, 2016). A study on adolescent exposure to tobacco and alcohol content in YouTube music videos by Cranwell et al. (2014) reported that a large number of British adolescents (especially women) view these video clips and are exposed to tobacco content. Similar results were found in a follow-up study on how adolescents are exposed to tobacco through visual tropes and song lyrics. Adolescents were exposed to around four times more tobacco content per year than adults (Cranwell, Opazo-Breton, & Britton, 2016).

Young people are the primary users of novel technologies like social media, and millions of them are making use of Facebook, Twitter and YouTube (Ballano, Uribe, & Munté-Ramos, 2014). Targeting young people with attractive (pro-tobacco) posts increases the reach for marketing smoking products, and in turn, young people may post smoking-related messages on their own social media pages. Research indicates that young people are keen to duplicate and disseminate material online if they consider it ‘cool’ (Cavazos-Rehg, Krauss, Grucza, & Bierut, 2014). This provides an effective way for tobacco companies to circumvent the ban on tobacco advertising. Maher et al. (2014) suggest that pro-tobacco organisations recognise this bias and should formulate social media messages accordingly as current online anti-tobacco messages are not attractive enough for young people to broadcast to their online network.
Moreover, social media provides a great deal of potential to advertise via the use of proxy advertisers (Elkin, Thomson, & Wilson, 2010), for example, having celebrities (proxies) post a picture on Instagram in which they smoke a cigarette. A particularly important form of proxy advertisement for young people are videos from vloggers (people who post videos on platforms such as YouTube) reviewing tobacco products (Elkin et al., 2010). Figure 3.2 illustrates an example of a tobacco review vlog which has been seen 58,789 times.

![Screenshot of a famous vlogger reviewing Marlboro Red on YouTube](image)

Figure 3.2 Screenshot of a famous vlogger reviewing Marlboro Red on YouTube (accessed on 28/09/2017).

Tobacco advertisement in films has been banned since 1990, but this does not include characters smoking unmarked cigarettes during scenes. So, even though the brand advertisement is gone, the promotion of smoking is still present. Two reviews on smoking in movies and youth smoking initiation by Charlesworth (2015) and Leonardi-Bee, Nderi & Britton (2016) discussed how smoking in films is an advertisement strategy to increase smoking uptake. Entertainment is a big part of young people’s lives and a great influence on their perceptions of ‘coolness’ and normality (Charlesworth, 2005). Although a direct causal relation cannot be made, the likelihood for people to initiate smoking doubles if they are exposed to smoking in films from when they are unexposed (Leonardi-Bee et al., 2016).

### 3.2 Tobacco control: helping people to quit

Tobacco control goes beyond the regulation of tobacco sales and advertisements to include programs and campaigns to help people quit smoking.
3.2.1 Anti-smoking campaigns

Every country that signed the WHO FCTC is required to direct attention to the annual ‘World No Tobacco Day’. This campaign is supplementary to the WHO FCTC (that encourages governments to implement tobacco control policies) and, by adding an annual day with a different focus each year, enables attention to be given to the wider issues surrounding tobacco control (WHO, 2015a). The day is an opportunity to promote new knowledge on the challenges of smoking from a global perspective. It can, for example, transcend local and national concerns by emphasising the environmental impact and growing global poverty that arises because of the tobacco industry and high rates of smoking prevalence (see Figure 3.3 for the 2017 poster).

Figure 3.3 Poster of the 2017 ‘World No Tobacco Day’ campaign (source: www.who.int/tobacco/wntd/ accessed on 28/09/2017)

As part of the WHO FCTC, partner nations are also encouraged to provide national campaigns to help smokers quit smoking and reiterate the dangers of smoking in everyday life through mass campaigning and local outputs (with both counselling and marketing). In the UK, public health organisations engage in both the annual ‘No smoking day’ and also a period of time where individuals are encouraged to quit smoking called ‘Stoptober’.
3.2.2 Youth-based initiatives

The national campaigns outlined above and mainstream smoking cessation support from organisations such as the National Health Service (NHS) tend to be aimed at adults. Youth-specific campaigns are mostly provided in school settings. Over the past three decades, school-based smoking campaigns have grown as it is the easiest way to influence youth smoking behaviour of even the most hard-to-reach young people (Mercken et al., 2012). These youth-dedicated programs need to be effective as pressure on school curriculum is increasing, and health education is not measured in school performance ratings (Langford et al., 2015).

Since the 1980s, school-based initiatives have evolved from passive information-only events into multimodal programs that challenge perceptions of smoking and enhance individual resistance to its attractive forces. Different approaches have been taken to achieve this, and initial results show an increase in intention-to-quit and higher risk awareness (Szatkowski et al., 2016; Vallone et al., 2015). Moreover, due to the internet, programs have progressed into campaigns with online access which makes it possible to engage young people in anti-tobacco interventions on a larger scale.

3.2.2.1 Multimodal programmes

According to a systematic review of the different types of anti-smoking programs in schools and their effect on youth smoking uptake by Thomas et al. (2013), the most effective programs focus on multiple facets of anti-smoking campaigning. They provide information, build confidence, convey life skills such as problem-solving and decision-making, and help develop cognitive skills for resisting interpersonal or media influences. The primary purpose of these interventions is to give young people the tools to quit smoking if they are smokers and the ability to resist peer pressure not to start smoking (R. E. Thomas, McLellan, & Perera, 2015). These programs vary between professionally-led interventions, peer-led interventions, and multimedia technology guiding young people.

Arguably, the most common approach is the professional-led multimodal type program. These can be given by the teacher or a stand-in professional. An example from the US is the American Lung Associations ‘Not On Tobacco’ (N-O-T) program that focuses on teaching young people the health consequences of smoking and training them to cope with emotions and how to avert peer-pressure. This program has an increased intent-to-quit rate of 15-19% (Branstetter, Horn, Dino, & Zhang, 2009). Another example is The Filter Wales workshop in which youth workers visit schools to teach young people about smoking
through various activities. More details about this program will be given later in section 3.4.

Another approach is to have peer educators lead the multimodal cessation programs. ASSIST (‘A Stop Smoking in Schools Trial’) is such a school-based ‘peer-led’ intervention. Influential students are asked to encourage other students in informal conversations not to smoke (Hollingworth et al., 2012; Mercken et al., 2012). Mercken et al. (2012) studied the effectiveness of the ASSIST program on adolescents with low and high socioeconomic position (SEP) in several European countries including England and Wales. Their results show how peer-led interventions are more effective than ‘professional’ ones and most effective in more deprived areas. Mercken et al. theorised that this difference occurs as young people with lower SEP tend to dismiss authority and are more likely to take advice from peers than similar students with higher SEP. The authors continue by expressing how peer-led interventions could reduce youth smoking inequalities more effectively than other types of tobacco control programs (Mercken et al., 2012). Campbell et al. (2008) studied the intervention effect of the ASSIST program in 59 schools from England and Wales. Their results revealed that the ASSIST program is most effective in rural and deprived areas, which are understood as the locations with the most hard-to-reach young (hardcore) smokers.

3.2.2.2 Beyond school interventions

Langford et al. (2015) carried out a systematic meta-analysis of school-based interventions and concluded that three elements were necessary for the success of the intervention: input into the curriculum, changes to the school’s ethos or environment, and engagement with families and local communities. This suggests that, while the school is an essential factor, the environment within which students live also has to be ‘on board’ for successful outcomes (Langford et al., 2015). Therefore, several youth-dedicated approaches reach beyond school interventions.

One such program is the European Smoking prevention Framework Approach (ESFA) which is designed to reach adolescents at four levels: the individual, parental, school, and community level (De Vries, Mudde, et al., 2003; Vartiainen et al., 2007). The ESFA programme is based on the Attitude, Social influence, self-Efficacy Model (ASE) as a theoretical framework, hypothesising that the key to a successful health behavioural intervention is the change in attitude and reduction in social influence on the individual (Vartiainen et al., 2007). The programme had a significant effect on the onset of weekly smoking but relied heavily on the effort made by the community and the available funds to
run this program for more extended periods of time in participating countries (De Vries, Mudde, et al., 2003).

The most internationally recognised and replicated program is ‘The Truth® Campaign’ from the USA, which was established in 1999. The Truth® campaign is an evidence-based, national smoking prevention campaign designed to reach at-risk youth (Allen, Vallone, Vargyas, & Healton, 2009). The program gives out accurate smoking information through (for young people) attractive campaigning and encourages them to become part of the movement to discourage smoking without preaching or stigmatisation (Allen et al., 2009; Evans, Wasserman, Bertolotti, & Martino, 2002). The campaign is mostly present online on social network sites such as Twitter and Instagram, and through advertisements on popular youth channels such as MTV. One of their ‘grand campaigns’ in 2016 showed ‘cute’ pets with the message that smoking kills the animals too (Figure 3.4). The focus on animals was wholly intentional, linking the branding of the campaign to the interests of adolescents. The philosophy of the campaign suggests that when the youth’s self-image and interests coincide or align with the focus and imagery used by the Truth® campaign, then they are more likely to take on the values promoted by the campaign (Evans et al., 2002).

![Figure 3.4 A still of a Truth® commercial about the increased chance of cancer for pets if the owner smokes (source: www.thetruth.com/ accessed on 28/09/2017).](image)

A large part of the Truth® programme is a counter-marketing campaign about the practices of the tobacco industry (Allen et al., 2009; Vallone et al., 2015). Truth® encourages young people to resist smoking influences by knowing how the tobacco industry operates. Research has shown that aggressive counter-marketing campaigns lead to more negative
attitudes and beliefs towards the tobacco industry than non-aggressive campaigns (Hersey et al., 2003; Richardson, Green, Xiao, Sokol, & Vallone, 2010). Awareness of the campaign was high among all adolescents and the intention to smoke decreased with higher awareness (Allen et al., 2009; Richardson et al., 2010).

An example of a cross-national transfer of the brand of the Truth® campaign is a UK anti-smoking school-based campaign, ‘Operation Smoke Storm’ created by Kick It, the UK NHS Stop Smoking Service. In this campaign, online programs provide school-based interventions focused on giving the truth about the tobacco industry. These programs move away from passive listening to actively engaging in investigative games (Szatkowski et al., 2016). Students (in year 7 and 8) act as secret agents to uncover the tactics of the tobacco industry as well as receive information on the risks of tobacco smoking. By using multimedia technology, this program can be implemented at a low cost and delivered by teachers, which makes the program highly accessible to schools in more remote or deprived areas. This is, especially in combination with the Truth® campaign, an impactful way of influencing young people to question if they want to contribute to the tobacco industry or join the campaign and expose its practices.

3.2.3 Online awareness and counselling

As shown above, many school-based programs utilise a multimedia approach, and many other organisations use online platforms for their messages as well. Nowadays, most organisations have an online presence via a website or at least an email address, and it has become more popular to have a presence on social network sites too. Health organisations advertise or present their campaign to the public through social network sites, and online health promotion has helped shape the health-related knowledge of a population (Chou, Prestin, Lyons, & Wen, 2013).

Nowadays, online smoking cessation counseling is done via social network sites (Ramo, Liu, & Prochaska, 2015; Struik & Baskerville, 2014), blogs (Brandt, Dalum, Skov-Ettrup, & Tolstrup, 2013), mobile apps (Paay et al., 2014), and alongside other programs with an online presence such as Stoptober. Shahab & McEwen (2009) did a systematic review of the literature on online support for smoking cessation and concluded that web-based cessation programs have the same potential as face-to-face counselling. Moreover, an advantage is that once the online program was set up, it could run for an extended period without much extra cost (Shahab & McEwen, 2009). Park & Drake (2015) add to that argument in their systematic review of online smoking cessation programs that online programs can be accessed at flexible times and locations and can have a tailored approach
just like face-to-face sessions. The extensive exposure to and familiarity of young people with social media are advantages for smoking prevention and cessation to reach hard-to-engage youth.

3.3 Challenges of existing approaches

Smoking cessation programs and interventions raise awareness of the health consequences of tobacco consumption and the exploitative nature of the tobacco industry. They have been influential in lowering the uptake of smoking among young people. However, a number of young people remain smokers, so this section outlines the adverse effects of anti-tobacco regulation and the challenges that arise from anti-smoking initiatives in relation to youth smoking prevalence.

3.3.1 Marginalisation through smoking legislation

Anti-tobacco legislation has transformed the way in which people engage in smoking activity. Smoking has shifted from a practice that was regularly undertaken in public places and considered socially acceptable to an undesirable behaviour that is condemned to designated areas. Several studies have focused on the effect of smoking policies on the smoking population and the marginalisation that comes forth from these (e.g. Alexander et al., 2010; Frohlich, Mykhalovskiy, Poland, Haines-Saah, & Johnson, 2012; G. Moore, Holliday, & Moore, 2008; Parry et al., 2010; Ritchie, Amos, & Martin, 2010a; Rooke et al., 2013).

Respondents in a study of the ban on smoking in public spaces in Canada complained that non-smokers had claimed all the nice spots and that no public spot is left to smoke without receiving judgement (Bell et al., 2010). Rooke et al. (2013) argued that since the ban on smoking in public places in the UK occurred, smoking has shifted from a relaxing and pleasurable activity to one that is considered as deviant. Both smokers and non-smokers now show ambivalent feelings towards smoking, so much so that smoking identity has become a vulnerable construct that needs care and constant consideration to be maintained (Rooke et al., 2013).

A study in Scotland on stigma and smoke-free public places legislation by Ritchie, Amos & Martin (2010a) discussed how even though there was little evidence of tangible discrimination towards smokers; the smokers felt stigmatised and experienced a loss of status. One of the visible and felt forms of distinction came through smokers having to separate themselves from others by smoking in designated areas, and this separation creates social exclusion (Frohlich et al., 2010). A later study by Frohlich et al. (2012)
found that the meaning and significance of smoking for young people reflected and reinforced social differentiation, attributing smoking to a lower social class. Similarly, a group of young Welsh people were interviewed about smoking stigma for the report by Parry et al. (2010), and it illustrated how young people stereotyped smokers into ‘disadvantaged’ and ‘irresponsible’.

To combat these feelings of marginalisation, smokers may turn to what is called ‘smoking islands’ which form a collective resistance to tobacco control (L. Thompson, Pearce, & Barnett, 2007). These places are locations where smoking remains a normal activity and are designed by smokers to produce a local culture of smoking (L. Thompson et al., 2007). These places reinforce continued smoking, and the seclusion protects against marginalisation (Barnett et al., 2016).

For smokers who are not part of these smoking safe havens; societal marginalisation becomes internalised to the extent that smokers may become self-stigmatising. The study by Ritchie, Amos and Martin (2010a) mentioned earlier found a form of self-stigmatisation in which smokers condemn their own behaviour as problematic. The internalised feeling of stigma can be evaded through placing oneself as a ‘considerate smoker’, someone who is aware of their surroundings when smoking (Phillips, Amos, Ritchie, Cunningham-Burley, & Martin, 2007; Ritchie et al., 2010a). This concept countered the feeling of stigma as ‘considerate’ behaviour was more socially acceptable and not worthy of social disapproval (Phillips et al., 2007). For some, the ban on smoking in public places has emphasised considerate smoking instead of quitting.

Relevant to note, many of the policies and programmes described earlier can not only stigmatise but can widen inequalities. Less socioeconomically advantaged groups are both more likely to smoke and less likely to engage with policies and programs (Hutcheson et al., 2008). Conversely, more advantaged people engage more intensively with anti-smoking policies and programs and are more successful in quitting (Kotz & West, 2009). In this way due to the anti-tobacco, inequalities increase (Hiscock, Pearce, Barnett, Moon, & Daley, 2009; Kuipers, Monshouwer, Van Laar, & Kunst, 2015; S. Thomas et al., 2008).

3.3.2 Circumventing legal sales of tobacco

The change in age restrictions for the purchase of tobacco is well-intended, but research has shown that a significant proportion of young people rely on proxy-sales, illicit tobacco, and ‘well-disposed’ shopkeepers. For example, a survey by ASH Wales on young smokers in 2010 revealed the different ways in which young smokers accessed cigarettes. The
results indicated that 80% obtained supplies via friends or family, 20% through the use of vending machines (which are now banned in the UK), 8% were obtaining illegal supplies from unmarked vans or sellers in private dwellings, and 7% reported the ability to buy single cigarettes from shop owners (ASH Wales, 2010).

With the increase in the legal age for tobacco sales, Scottish 16 and 17-year-olds found it increasingly difficult to purchase cigarettes themselves and relied more on their social connections such as family and friends to purchase the tobacco for them instead (Borland & Amos, 2009). Obtaining cigarettes via friends and family is indeed a convenient way for young people to obtain tobacco, as these are also the people they smoke with (Hoving et al., 2007). Donaghy et al. (2013) note that while this proxy purchasing takes more time to access tobacco supplies, the success rate is higher than trying to buy cigarettes themselves, and is therefore worthwhile pursuing.

The survey by ASH Wales noted above, revealed that 8% of young smokers avoid smoking regulations through purchasing illegal tobacco. As the price of smoking goes up, smokers resort to illegal and smuggled cigarettes (L. F. Stead & Lancaster, 2000). As a consequence, raising taxes arguably results in increased use of ‘fag’ or ‘tab’ houses (Gough et al., 2013). M. Stead et al. (2013) undertook focus groups with people living in disadvantaged communities and discussed how fag houses had become standard features in those communities. Many smokers and ex-smokers mentioned going to these fag houses on a regular basis for their smoking supplies. Moreover, buying from these houses has become a norm instead of an effect of marginal behaviour; buying tobacco in a shop would be ‘showing off’ (M. Stead et al., 2013). M. Stead et al. continue by illustrating how these fag houses were particularly crucial for under-age smokers as it was cheap and there were no regulations. While smokers acknowledged the illegality of this form of tobacco sourcing and recognised a drop in tobacco quality, it is the desire to satisfy nicotine dependency which defeated all other concerns (M. Stead et al., 2013). However, there is also evidence from young people that such sources are avoided because cigarettes taste horrible and are unpleasant to smoke (Robinson & Amos, 2010). Furthermore, illicit or black-market cigarettes were not preferred and were consumed only as a ‘last resort’ (Donaghy et al., 2013).

Seven percent of the respondents in the ASH Wales study mention buying cigarettes from shop owners. Robinson & Amos (2010) studied tobacco access for minors in small communities and reported that it was likely that the local shopkeepers provide young smokers with tobacco. The young people from their study explained how they had
developed a strategy in which they only purchase at certain shopkeepers that will sell to them (Robinson & Amos, 2010). The young smokers have found others ways of getting their tobacco cigarettes and continue (or start) their smoking behaviour.

### 3.3.3 Challenges with smoking cessation campaigns

The literature is mixed regarding the success of anti-smoking initiatives centred on smoking cessation interventions. Grimshaw & Stanton (2006) concluded in their systematic review of these interventions that there is not sufficient evidence for an overarching model that works. One of the most significant challenges is that young people tend not to engage formally in tobacco control programs and do not use an official campaign or organisation to help them with their quitting attempt (Whittaker et al., 2010). This may be because many anti-smoking campaigns and organisations have a strong focus on helping smokers who are predominantly adults quit smoking (Brown et al., 2014). Furthermore, as also mentioned in the previous chapter, young people tend to make a quitting attempt without much planning and forget about it (Berg et al., 2010). However, some studies did find that quitting attempts coincided with highly promoted initiatives such as Stoptober (e.g. Vardavas, Filippidis, & Agaku, 2014; Wadley et al., 2014).

As the vast majority of smokers have their first cigarette while they are still in school, school-based anti-tobacco programs have the most significant possibility of influencing individual behaviour and the capacity to steer young people’s attitudes towards unfavourable opinions of smoking (Mercken et al., 2009). Indeed, such interventions (e.g. the Truth® campaign, as outlined above) have been shown to be effective, in the short term, in changing perceptions and providing young people with the tools to resist social influences (Allen et al., 2009; Vallone et al., 2015). However, a systematic review by Thomas et al. (2013) reported that the influence of such multimodal school-interventions (incorporating various forms of interactions) wears off within a year to the level that there is no visible effect on the youth that participates in the programs compared to the youth in the control group. To exemplify further, the ASSIST-program has been shown to reduce smoking uptake after one year, but there is no significant parallel increase in smoking cessation (Audrey, Holliday, & Campbell, 2006). Another example comes from Szatkowski et al. (2016). Their evaluation of ‘Operation Smoke Storm’ concluded that, even though the initial reaction seemed promising, it did not appear to have decreased smoking uptake or susceptibility one year down the line.
This literature review has highlighted important tobacco control initiatives and the accompanying challenges. It has been suggested that young people are more susceptible to peer-led initiatives, that they engage extensively with social media, and that they are less impacted by generic initiatives without a youth focus. These suggestions point to the need for a program like The Filter Wales. The Filter Wales is based on the previous anti-smoking campaigns, is extra school curricular, and applies advanced social media engagement to help young people. Their setting is Wales (a country with around 3 million inhabitants) which makes the playing field smaller and offers room for experimentation. This study focuses on this youth dedicated smoking cessation campaign in particular as the Filter provides an excellent research opportunity to examine what works in terms of persuading young people not to smoke or to quit smoking.

3.4 The Filter Wales

The Filter Wales, which was set up in 2013, takes a multimodal approach towards engaging with young people about the use of tobacco (and to a lesser extent other smoking products). It works largely through deliberate and targeted interaction via social media alongside traditional outreach events. The following paragraphs describe The Filter Wales in detail, outlining its ASH Wales origins and describing its specific elements. Although there has been little in-depth research of The Filter Wales, there has been an evaluation by Meek, Hurt & Grant in 2015. The section, therefore, concludes with a summary of this evaluation, highlighting the gaps that this thesis will address.

3.4.1 Background of The Filter Wales

Action for Smoking and Health (ASH) was set up in 1971 as a public health charitable organisation that campaigns to eliminate the harm caused by tobacco smoking through judgement-free tobacco control activities. The two main agenda points of ASH are ‘information and networking’, meaning spreading smoking-related facts and increasing awareness of the danger of smoking, and ‘advocacy and campaigning’, signifying challenging the government on policy to reduce the harm from tobacco use (ash.org.uk/about). ASH Wales is the Welsh branch of the organisation, and The Filter Wales is a division, specific to ASH Wales, which focuses on smoking and young people. Similar to ASH more generally, ASH Wales raises awareness of the dangers of smoking and works together with communities and partners across the country for a smoke-free Wales (Welsh Assembly Government, 2011). A screenshot of their website can be seen in Figure 3.5.
ASH Wales received a three-year grant (£864,881) from Big Lottery People and Places Grant and a Big Lottery Innovation Grant (from 2013 to 2016) to fund a youth-dedicated smoking service providing smoking-cessation support and smoking prevention to 11-25 years-olds in Wales. This award provided the funding base for The Filter Wales as illustrated in the screenshot of the website (Figure 3.6).
The Filter is our youth service which provides quit smoking support and smoking prevention to 11-25 years olds. The Filter’s motto is to “Filter out the myths and give the facts about smoking” to young people.

The Filter’s services along with its exceptionally experienced youth workers are now available for commissioning.

Alongside the basic health impacts of smoking and tobacco, the Filter delves into related topics to spread the facts about:

- E-cigarettes
- Shisha/hookah
- Illegal tobacco
- Co-consumption with cannabis and legal highs

What we do

The project specifically targets those classed as ‘hard-to-reach’ in a number of ways:

- Workshops in a variety of youth provisions including: youth centres, Pupil Referral Units, young parent groups
- Workshops specifically designed for schools
- Two specific, highly targeted projects: Commit to Quit and Filter Apprentice
- Comprehensive social media and online engagement via the most prominent channels including Facebook, Twitter and Snapchat
- An informative website which provides support, and the facts about tobacco, in a relevant and targeted way for young people
- Volunteering opportunities in a number of aspects of our work

Find out more

The Filter website contains a wealth of youth-focused information about smoking and tobacco, alongside the topics mentioned above. It also includes a comprehensive section for professionals who work with young people which explains more about our various strands of work and has lots of free resources for youth provision.

Commissioning

The Filter provides an innovative and multi-strand approach to addressing smoking and tobacco issues with young people. The service is flexible and can be adapted to a range of professional disciplines, youth settings and groups of young people. Our website and social media channels complement our youth work and training services and we offer a cost-effective solution which can be tailored to your needs.
3.4.2 The Filter Wales campaign

The Filter’s aims are, besides smoking-cessation support and smoking prevention, to present the facts about smoking and to filter out the myths, e.g. the dangers of second-hand smoking and the content of tobacco cigarettes. Their approach of presenting non-judgemental facts is comparable to that of the Truth® campaign in the United States and is similarly branded to appeal to young people. However, the Truth® campaign has a strong emphasis on aggressive counter-marketing against tobacco companies (Richardson et al., 2010) whereas The Filter Wales has a less contentious approach. The outreach events offered by The Filter are specifically designed for young people and engage them in thinking about smoking issues through various activities that can be adapted to requirements.

The youth development team worked with over 5500 young people face-to-face at more than 220 sessions since the project’s launch in 2013 (see the website of Figure 3.6). These workshops are generally given in locations where young people come together such as schools, universities, youth centres and Pupil Referral Units.

Usually, organisations such as youth clubs request The Filter Wales to deliver workshops at which the Filter youth workers spend one session per week with the same group for six weeks. They present different smoking-related projects each week alongside the regular ‘Commit to quit’ focus. The ‘Commit to quit’ program entails young people blowing into a carbon monoxide (CO)-monitor to measure the carbon monoxide in their lungs from smoking (in the last 24 hours). Every week the CO-levels in their lungs are measured and the people with the largest decline across the weeks win a prize (generally a store voucher). A typical parallel activity is a competition for the best ‘no smoking’ advert led by ‘Cut Films’, a nation-wide initiative supported by The Filter Wales. This advert can be made during the workshops with the help of the youth workers, or independently by a school or youth club.

Besides the outreach events, the Filter uses the most prominent social media channels including online accounts on Facebook, Twitter, and Snapchat to create a comprehensive social media and online engagement. A social media team works on the Twitter and Facebook pages, Snapchat, and The Filter Wales website, posting information about the campaign, programs and competitions.
3.4.3 Evaluation of The Filter Wales

An evaluation of The Filter Wales was commissioned in 2015 by ASH Wales and was undertaken by Meek, Hurt & Grant of the Cardiff University’s Institute of Primary Care and Public Health. This evaluation covered the entire Filter Wales project and involved semi-structured interviews with the Filter staff, professionals trained by the Filter, and young people that were in contact with The Filter Wales program. A survey was created to learn from the professionals and young people about their experiences of all elements of the program. The evaluation included a thematic analysis of the tweets from The Filter Wales Twitter page and an assessment of young people’s awareness of other social media, i.e. the website and Facebook.

The results of the evaluation showed that the flexibility of the Filter program helped the team to adjust their plans to the specific circumstances of different workshop groups as well as to the desires of schools or youth clubs. This flexibility also enabled the Filter to remain up-to-date and connected with their audience. All parties (i.e. the Filter team, young people, and professionals) regarded the workshop program very positively. The range of activities and the interactive nature of the workshops kept young people interested, and with the help of visual aids, the non-smoking message was well received by young people. The use of social media and the website enhanced the delivery of information and enabled informal interaction with young people. However, according to the evaluation, it was mostly professionals who engaged on the social network sites instead of young people (Meek, Hurt, & Grant, 2015).

3.4.3.1 Evaluation of The Filter Wales Twitter feed

A specific part of the evaluation consisted of examining The Filter Wales Twitter feed. The collection period for the evaluation of the Twitter element was 35 weeks of data from 2013 and 33 weeks from 2014 which provided 1816 tweets from that period but did not cover Twitter usage consistently over the two years.

The researchers evaluated the content of the tweets and identified five categories: response to a third party (21%), an original tweet (27.6%), a retweet of a third party (51%), tweets sent to the Filter (6.8%), and unclassified tweets (0.4%). The original tweets from the Filter covered a wide range of subjects, for example, prize winners and outreach workshops, and their ‘replies’ usually contained an encouragement to the young people to quit smoking.

The content of the tweets was categorised by the following themes (which were not mutually exclusive): The Filter resources (e.g. posters), tobacco event promotion (e.g.
Stoptober), information sharing, tobacco control campaigns, quitting smoking, and unclassifiable. The evaluation gave an insight into the Twitter feed of The Filter Wales and illustrated the efforts made by the Filter social media team, but the scope for more extensive analysis is considerable.

3.4.3.2 Critiquing the evaluation of The Filter Wales

The overall evaluation by Meek, Grant & Hurt (2015) revealed that engagement with ‘hard-to-reach’ young people in Wales is difficult to achieve mainly because the workshops and social media engagement are voluntary. It focused on the services provided by The Filter Wales and did not go into detail about the young people that engage with the program. The individuals participating in the evaluation did not respond in high numbers which made it difficult to generalise results nor did the evaluation look into the deprivation context or rural/urban differences. This is a missed opportunity as it is known (and has been shown in the literature review) that deprivation levels and rurality are important factors in smoking prevalence, successful quitting smoking, and perceptions of marginalisation.

The evaluation of The Filter Wales Twitter feed was incomplete in its data collection, and the content analysis was superficial with respect to the type and topic of the tweet. By focusing the analysis purely on what The Filter Wales was tweeting, the evaluation failed to review and summarise the range of content that young people tweeted in the interaction with the Filter Wales. More research needs to be undertaken in terms of focusing on the young people’s exchanges with the Twitter feed, and a more in-depth analysis of the Twitter activity of these young people has the potential to shed some light on how they perceive and engage in smoking and other health risk behaviours. This involves a more detailed breakdown of the smoking-related themes captured in the tweets, retweets, and the more comprehensive tweet exchanges made by this youth group.

3.5 Aims and concluding remarks

This chapter has pointed out the limitations of existing approaches to decreasing youth smoking and shown where interventions lack effectiveness. As outlined in Chapter 1, while smoking prevalence is slowly reducing, the supply of young smokers into the smoking pool has not yet halted. The different anti-tobacco approaches outlined in this chapter highlight the complexity of approaches that are needed to address this critical public health issue and a particularly innovative approach is possibly offered by The Filter Wales. It has placed The Filter Wales in context, noting its antecedents and assessing the limitations of
the existing evaluation. Importantly, the chapter has indicated that the young people’s Twitter encounters with The Filter Wales provide a unique, precious opportunity to understand more about what smoking activity means for a large sample of young people and the extent to which it plays a part in their everyday lives. Such information is crucial to understand the social practice of smoking and to design interventions that embrace this nuanced understanding of this health-damaging activity.

This study explores the reach of the Filter by examining the spatial and temporal dimension of the Twitter activity. It also looks at the type of places from which activity originates in terms of levels of social deprivation and levels of urbanicity. The first aim is;

1. to evaluate the reach of the Twitter element of The Filter Wales campaign.

This study also gives detailed attention to the content of the Twitter feeds of the young people who have interacted with The Filter Wales. The Twitter content is unravelled to find out more about the social meaning of youth smoking. Part of that understanding includes a perspective of smoking in relation to other health risk behaviours detailed in Chapter 2 section 2.6. Much of the research literature, reviewed in this and the previous chapters, used traditional methods such as questionnaires, interviews, and focus groups to uncover the many facets of youth smoking. Participants were aware of their contribution and were required to identify their perceptions of smoking behaviour. This leaves a gap for novel approaches to add to knowledge by identifying the aspects of smoking that are central for young people without the interference of a researcher. Moreover, the research presented here exploits a unique opportunity to scrutinise a large volume of qualitative material within a ‘big data’ framework. Although the research data presented here is relatively short across each tweet, the snippets of text originate from thousands of tweets and in this sense strengthen the generalisability and reliability of the study. Consequently, the second aim of this research is;

2. to analyse the text content of tweets about smoking to understand more about the social meaning of smoking and other health risk co-behaviours.

The two aims are broken down into six objectives. The first two objectives relate mostly to the first aim to uncover the reach of The Filter Wales. At the end of this study there will be clear evidence about:

1. gender and age differences in the use of Twitter.
2. the geolocation of the Twitter users.
The next four objectives are linked to the aim of better understanding the social meaning of smoking. At the end of the thesis, more is known about:

3. differences in the content of tweets.
4. variations over time and place concerning the sentiments evident in smoking-related tweets.
5. the extent to which co-behaviours are present in the smoking-related tweets.
6. the wider context of smoking for young people as evidenced by their Twitter archive.

This chapter has stressed that a full evaluation of the geographical and socio-demographic reach of the Filter is needed to assess whether it is likely to engage with those parts of the youth demographic who may be harder-to-reach through conventional approaches. Importantly, the Twitter exchanges generated as part of this intervention provide a unique opportunity to have a fuller insight into the social meaning of smoking outside of a formal research framework. Findings are likely to help public health workers understand more about why and how young people persistently engage in this health-damaging behaviour.

The following chapter outlines the methods used to convert the tweets associated with The Filter Wales into research data for this thesis after briefly outlining how Twitter data can be used in social science research. Chapter 5 then focuses on the methods used to address the two broad aims and the specific objectives listed above.
Chapter 4. The research use of Twitter data

This chapter is the first of the methods chapters and outlines how Twitter can be applied as a research tool and used in fulfilling the aims and objectives of this thesis (stated in the previous chapter) by outlining how the data is collected and prepared. This whole chapter explains how Twitter-related information forms data for this thesis.

The first part of this chapter discusses what Twitter is (section 4.1), elucidates how this form of social media can be regarded as a novel source of research data (section 4.2) and provides an assessment of the ethics of undertaking research involving Twitter data (section 4.3). The second half of the chapter focuses on how the Twitter data was gained and used as research data within this study. The first section here (4.4) explores how and what Twitter data were gathered, and how they were prepared and categorized for analysis within this thesis (in section 4.5). The last section (4.6) outlines the meaning of text-elements and retweets in the Twitter data.

4.1 What is Twitter

Twitter is a social network site created in 2007 based on the activity of microblogging (i.e. posting small pieces of text). Twitter has become one of the main social network sites in the world with over 328 million active users and 500 million tweets sent out every day in 2017.² The UK has approximately 16.5 million Twitter users in 2017 which is approximately 25% of its population.² This social network site is also used by numerous companies and organisations to promote their activities and products, but most of Twitter accounts belong to individuals (Cover, 2012). For these people, Twitter is used for information gathering, reading or sharing experiences and opinions, maintaining social relationships, and to a lesser extent engaging with others on socially relevant matters (Cover, 2012; P. R. Johnson & Yang, 2009; Marwick & Mand, 2010).

To give an impression of who uses Twitter, a website called ‘Think digital first’³ describes the demographics of Twitter users in the UK in 2016. The website states that there is an almost equal gender divide (51% female and 49% male) and 65% of these Twitter users are under 35.³ A study using the British Social Attitudes Survey 2015 to uncover Twitter usage showed that people with managerial, administrative and professional occupations were almost twice as likely to have a Twitter account (Sloan, 2017). However, this study

³ https://www.thinkdigitalfirst.com/2016/01/04/the-demographics-of-social-media-users-in-2016/ accessed on 25/05/18
did not combine locational data to Twitter use. A study on social media use in the USA did measure where the Twitter users were from and revealed that two-thirds of them were from urban areas (Duggan, 2015). These demographics signify that Twitter is mostly used by the younger and higher educated people predominately living in urban areas.

4.1.1 What is a ‘tweet’

The microblogs placed on Twitter are called tweets. A tweet is a string of words consisting of a maximum of 140 characters\(^4\). As an addition to words, tweets can contain an ‘@’-sign, hashtag (#), an extension (i.e. the third party developed applications), and emojis and short clips (e.g. GIFs; see below for further elaboration). All these elements are used to provide context to the content and show the intention of the originator within the character limit.

The ‘@’-sign is used to link the tweet to another Twitter user, and this individual will receive a message that their Twitter name came up in someone else’s tweet. For example, the following tweet was made by a lecturer in Health Geography when she saw me at a conference:

"@Kim24501 last but not least at #imgs2017 really interesting talk on twitter and smoking."

The hashtag serves a different and more complex purpose as it can connect a tweet to a wider context such as the above example linking the tweet to the conference (#imgs2017) or relate to ‘trending topics’ such as #GameofThrones or #Brexit. A trending topic is a subject that is tweeted about many times within a short timeframe (Tsur & Rappoport, 2012). The Twitter user can click on the hashtag and get all the tweets which include that specific hashtag on one page. This is convenient if the individual wants to have all the tweets about that trending topic.

Tweets can also have an extension. These links can be added to create easy access to a website, a video, or a picture which would otherwise exceed the character limit if written out fully. Two other possible extensions are emojis and Graphic Interchange Formats (GIFs), but they are used differently than the other extensions. Emojis and GIFs are shortcuts to express non-verbal cues of the originator. Emoji is a broader term which means ‘picture character’ and emojis are used on a regular basis to let the receiver of the message know about the intention behind the tweet. Most emojis represent depiction that cannot easily be expressed through written words. Examples of these emojis are ‘happy, ☺

---

\(^4\) This has been increased to 280 characters in November 2017.
or :)’ and ‘sad, 😞 or :(’. A relatively new feature of Twitter is the possibility to add a GIFs file. These GIFs are small clips that (similar to emojis) add a visual element to the tweet. All the items mentioned above add context and additional information to the words in the tweets and together make up a tweet. This context is important for the interactions with other Twitter users.

4.1.2 Followers on Twitter

Twitter is a social network site in which the users choose whom they would like to ‘follow’ on Twitter. The user can follow anyone without knowing them offline, and as a result, many celebrities and news-related pages have numerous followers (e.g. Bill Gates has 33 million followers and only follows 183 others). By following someone, people receive all of the posts created by the people they follow on their Twitter home screen. This creates a Twitter network of followers with similar values and interests, viewing and responding to tweets of others and provides an opportunity to connect to the world (M. S. Smith & Giraud-Carrier, 2010). However, people mostly interact with others they know offline too and use Twitter to strengthen their bond (Subrahmanyam, Reich, Waechter, & Espinoza, 2008).

From a social perspective, in comparison to other social network sites such as Facebook, Twitter has the lowest level of direct engagement between individual users (D. J. Hughes et al., 2012). Here direct engagement is defined as any type of interaction through a reply, a ‘like’ (i.e. make notice that they appreciate the tweet), and a ‘retweet’ (i.e. the tweet is copied onto the account of the individual). With regard to Twitter posts, only a fraction of the people that can access the tweet will interact in any form.

Twitter gives users the possibility to remain anonymous as the personal information in the Twitter profile is limited and, in theory, can be completely fictional. The idea behind fictional Twitter profiles is that anonymity creates a new perspective on posting online as it encourages more honest posts and responses, according to Hughes et al. (2012). However, this facelessness of Twitter also increases the number of ‘trolls,’ i.e. people who use the social network site to provoke others and disrupt conversations by posting extraneous or derogatory messages (Lampe & Johnston, 2005).

Overall, tweets are deliberate expressions to connect with others who (might) find the offered information interesting. The topics covered in the tweets represent important themes for these people and are central components for the understanding of meaning that people transmit online.
4.2 Twitter as a research tool

This study illustrates how Twitter can be used as a tool, or more specifically, a form of relevant data in smoking research. Twitter offers a substantial source of data that is widely used in research by corporations, organisations and academia. For industry, Twitter information is used to help them develop their marketing strategy and, for example, may be used to research how to better influence nicotine users into buying their products and to advertise their products more efficiently (Kavuluru & Sabbir, 2016). Organisations (specifically for this study; health organisations) use Twitter as a research tool (occasionally) to discover if and who received their information, e.g. Harris et al., (2013) studied the network of local health departments in the US and uncovered that they mainly disperse information amongst each other. Similar, The Filter Wales team found that their information through social media is primarily picked up by other health organisations and professionals. Lastly, academia has focused mainly on using Twitter for uncovering patterns in human behaviour or, when using it with a health focus in mind, on the spread of ill-health. This section goes deeper into the ways in which Twitter is used as a research tool in academia.

4.2.1 Examples of Twitter as an academic research tool

A significant amount of research has been done on examining the transmission of health information through posts placed on Twitter (Culotta, 2010; Signorini, Segre, & Polgreen, 2011). Within health studies, research mostly consists of empirical analyses of Twitter posts by prevalence and disease spreading through time and space on subjects such as the flu, concussions, insomnia, toothaches, problem drinking, and obesity (Ghosh & Guha, 2013; Heavilin, Gerbert, Page, & Gibbs, 2011; Jamison-Powell, Linehan, Daley, Garbett, & Lawson, 2012; Sullivan et al., 2012; J. H. West et al., 2012). These empirical studies share common methods; during a defined period (which varies considerably among the different studies), the researchers collected data on their specific topic by using related search terms. Subsequently, content analyses are carried out to denote how the disease is talked about, and this can reveal time and space patterns of the health issue on this platform.

For example, Sullivan et al. (2012) studied mentions of concussions in tweets through the search terms ‘#concussion’, ‘#concussions’, ‘#concuss’ and ‘#concussed’. All the tweets that came up through the Twitter collection program (which will be discussed in Chapter 4 section 4.4.2) were analysed in terms of their textual content. The researchers concluded
that Twitter could be used to reveal ‘snapshots’ of everyday health as well as testing the awareness and accurate information about concussion from the Twitter users (Sullivan et al., 2012). They found that one-third of the concussion-related tweets came from health organisations (33%), followed by ‘personal updates’ from individuals (26.8%). The authors theorised that the personal updates are used as thought outlets and can help people relieve emotional tensions. Similarly, Heaivilin et al. (2011) studied dental pain through set search terms, coded them regarding their content, and argued that Twitter presented a unique opportunity to uncover how people are experiencing dental pain. These content analyses studies argue that Twitter can be utilised as a tool to uncover experiences, attitudes and beliefs about health topics in everyday life (Heaivilin et al., 2011). This group of tweets enhanced the understanding of the layman's perceptions of feeling healthy and the meaning of this illness (Heaivilin et al., 2011; Sullivan et al., 2012).

The following paragraphs display some of the different terrains smoking studies have covered besides content analysis. The first example is smoking progress through the Twitter histories of the sample population. Among others, Murnane & Counts (2014) in their study on the progress in smoking cessation covered on Twitter among 653 adults argued that Twitter can uncover patterns of success and failure in quitting smoking attempts. They discussed this through a search for ‘quitting smoking’ related terms, and when the quitting smoking users were identified, they collected more tweets from the same users. With this information, Murnane & Counts (2014) identified the quitters and relapsers and determined where they differed in their Twitter use, e.g. relapsers tweeted more at night and quitters had more ‘goal setting’ tweets. The aim of the study was to uncover clues in Twitter content that could indicate beforehand who would be successful at quitting smoking (Murnane & Counts, 2014).

Another area of Twitter-related smoking research focuses on the positive and negative connotations of smoking within smoking-related tweets. Twitter studies provide a unique insight into opinions and sentiment of smoking (Cavazos-Rehg et al., 2015; Krauss et al., 2015; Myslín, Zhu, Chapman, & Conway, 2013). Myslín et al. (2013) collected 7326 tobacco-related tweets and classified them by content and sentiment. They argue that people use more positive sentiment towards smoking on Twitter than they were likely to do in surveys, especially concerning newer products like e-cigarettes and shisha. As most of the posts are countenances of emotions or opinions, it is likely that Twitter reveals the unadjusted sentiment towards smoking and tobacco use (Myslín et al., 2013). Harris et al. (2014) researched Twitter comments of the general public on a health campaign for e-
cigarettes by the Chicago Department of Public Health. Their results show that tweeting against the campaign was ten times greater than tweeting in favour of the campaign and the majority of the negative tweets came from private Twitter users as opposed to health professionals and policymakers who were most positive (Harris et al., 2014).

There are many other studies focusing on smoking on Twitter with a different aim such as focusing on smoking Twitter profiles (e.g. Prochaska, Pechmann, Kim, & Leonhardt, 2013), focusing on other smoking products such as marijuana and shisha (e.g. Cavazos-Rehg, Krauss, Grucza, & Bierut, 2014; Krauss et al., 2015), focusing on the smoking social networks (e.g. Rocheleau et al., 2015), and using the example of tobacco to showcase models for tackling public health problems (e.g. Prier, Smith, Giraud-carrier, & Hanson, 2011).

This study follows similar methods to these examples. However, none of the previous studies has done exactly the same as what is being done here. There are elements of many of them (such as the content analysis, the view on process and revealing patterns of time, space and sentiment analysis) integrated into this study. This study combines these methods to illustrate a larger picture of the Twitter data and the possibilities with a Twitter-based smoking-related dataset. The following sections reveal what advantages and challenges come with utilizing this type of data.

4.2.2 Challenges of traditional methods compared to Twitter research

For decades, research on smoking and youth has been done via interviews, focus groups and surveys. These research methods have provided interesting insights into the prevalence, the context, and meaning of youth smoking. The challenges of using traditional methods have been touched upon in the introduction chapter and literature reviews and showed how quantitative studies commonly miss depth and qualitative studies rely on low sample numbers. Moreover, there are a few common disadvantages to these approaches such as participants trying to fit the norm and not providing truthful responses which may lead to underestimation of smoking. There are also challenges concerning the validity and reliability of retrospective assessments. The next few paragraphs elaborate on these and how using Twitter data avoids that challenge.

In many studies into young people’s smoking and other health-related behaviours, the participants were aware of their participation in the study, and they are more likely to embellish the truth when they are directly questioned about the activity (Myslín et al., 2013). Twitter does not have a governing body scanning for what is placed on Twitter
(which does exist for other social media sites such as Facebook) meaning that there is no censorship of posts and profiles. This is an advantage of using Twitter as the data is uncorrupted and does not (necessarily) conform to social norms and appropriate social behaviour. Tweets can contain highly offensive comments.

More traditional data collections approaches are centred around a specific topic. As Twitter research is data-driven, the participants are not required to express their thoughts on that topic. Instead, they post content that resonates with the followers of their Twitter network (M. S. Smith & Giraud-Carrier, 2010). The data-driven information presents a glimpse of their thoughts without the interference of a researcher. The place of the topic (in this smoking) in everyday life appears by analysing Twitter data that is created without this insistence.

Multiple studies highlighted that respondents in smoking-related studies tend to underestimate their smoking behaviour (e.g. Berg et al., 2009; Heikkinen et al., 2010; Leas et al., 2014; Moodie et al., 2008). For example, people often do not consider themselves smokers because they compared themselves to others that smoke more or because of the negative association with that label. Twitter, on the other hand, is data-driven, and the content of the tweets discloses their smoking status.

Challenges in retrospective assessment in smoking research are exemplified by Mair et al. (2006). Mair et al. demonstrated that the self-reported smoking accounts, based on the smoking accounts of youth in the UK Liverpool Longitudinal Smoking Study (LLSS) over several years, were asymmetrical. When questioned about their first experience smoking a cigarette, the respondents’ stories changed each year. Among other things, they found that especially in retrospect many mistakes are made. The self-assessment of the young sample members were subjective and vulnerable to changes in meaning through the years (Mair et al., 2006). A change in their meaning of smoking is not a challenge in Twitter research as the Twitter feeds are not self-assessments but self-expressions from the individual at specific moments in time.

Twitter has become a popular research tool as it has several advantages over traditional research tools. Often, high correlations are reported between Twitter statistics and real-world statistics. For example, similar results between flu-related admissions to health services and flu-referenced Twitter collection are found from two studies (Culotta, 2010; Signorini et al., 2011). The advantage of Twitter is that the data is more extensive and faster than health service reports, making Twitter a rapid, cost-effective health status surveillance (Myslín et al., 2013).
4.2.3 Challenges of using Twitter as a research tool

Clearly, there are possible advantages to using the microblogs of Twitter to get a better understanding of smoking behaviour. However, there are also a number of possible disadvantages to the use of these posts in youth smoking-related research.

One of the key issues which must be addressed is the possibility of misrepresentation, i.e. when the reported activity contained in a tweet does not necessarily coincide with real-life. For example, people use Twitter to share information they think their followers appreciate and post personal updates as cathartic experiences (Jamison-Powell et al., 2012; Quercia, Capra, & Crowcroft, 2012; Sullivan et al., 2012). Moreover, the posts might not be truthful, and there is no operative and reliable way to evaluate the trustworthiness of Twitter feeds on a larger scale (Bryman, 2012). Individuals often use Twitter to show a favourable version or view of themselves to their followers. The mentioning of smoking or any aspect of people’s lives does not necessarily reflect ‘real life’ and could be posted as a way to ‘look cool’ (Moreno, Briner, Williams, Walker, & Christakis, 2009).

As previously mentioned, Twitter can be used to reveal spatial-temporal patterns of poor health, sometimes in a more timely manner than health-service related disease and illness monitoring. However, it can lead to ‘false positives’ or even ‘false negatives’ whereby a pattern is found which is more associated with the rate of Twitter activity than with the health issue itself. Twitter exposure is not always equal to the cases of the epidemic and Twitter is highly reliant on the popularity of the moment as exemplified by a study on the bird flu. Signorini, Segre & Polgreen (2011) showed that while the number of tweets on the H1N1 bird flu virus was going down, the actual number of cases was going up (a false negative). Moreover, in a study on obesity in the United Stated, Ghosh & Guha (2013) showed that people in the bigger cities on the East and West coast tweeted more about obesity even though their obesity rates were not as high as in the Midwest (a false positive). These tweets do not coincide with the prevalence of the illness but with Twitter popularity.

The advantages and challenges of using Twitter as a research tool relate to the features of Twitter itself. The data is uncensored and presents an opportunity to uncover information that is difficult to gain otherwise. But, this data can misrepresent sample members as Twitter activity embodies a version of reality people want to portray online.
4.3 Ethical considerations for Twitter research

In theory, the use of Twitter holds no ethical concerns. Tweets are publicly available, and the Twitter Company does not limit to the gathering of these tweets. Many studies refrain from mentioning ethics as the Twitter Company has a built-in application to get consent for the use of the individuals’ feed by third parties. People signing up for Twitter need to agree to the terms and conditions presented which grant third-party public access (www.twitter.com/tos accessed on 16/11/2017). By accepting the terms and conditions, anyone can access and use the data provided by the individual on the Twitter social network site.

Not many Twitter studies are concerned about ethics as Zimmer & Proferes (2014) reported in a systematic review of 531 articles explicitly analysing Twitter data. They only discovered 16 studies mentioning ethical considerations, and out of those, only 5 (less than 1% of the total) acknowledged ethical concerns shaping the data collection and presentation (Zimmer & Proferes, 2014). Specifically, if people do not want their tweets to be public, this can be altered if the user changes his or her Twitter setting into a profile with ‘protected tweets’ in which only followers can see and access the tweets. This way Twitter users can safeguard their tweets from being accessed by third parties.

However, Twitter users are generally unaware of these specific terms and using Twitter data can be particularly sensitive as people want others to see their posts but not to the extent that it can be used in a possibly harmful context such as by insurance companies (Crawford & Finn, 2015). There is a dilemma that occurs as the idea of self-management of data and privacy is rarely achieved (Solove, 2013). It is difficult for individuals to assess if any piece of information will be used and with the combination of other data reveal something sensitive (Solove, 2013). Posts on Twitter could identify individuals, for example, advances in data linkage have in combination with the message content increase the possibility for the exploitation of time/date and geographical location to build a detailed picture of people’s lives (McKee, 2013). Research into privacy leaks through Twitter posts has shown that a significant portion of tweets included personal information about the author and posed potential privacy threats (Mao, Shuai, & Kapadia, 2011). Moreover, Moa et al. (2011) found that Twitter users were often unaware of the fully public nature of their Twitter engagement and its possible abuse by third parties.
4.3.1 Following ethical guidelines

The use of Twitter data in the study is to understand smoking from the perspective of young people better and serves no other purpose than uncovering their social meaning and the possible social inequalities in this sample population. Several papers discuss the potential for social media (such as Twitter) to be damaging to the individual and present guidelines to overcome such difficulties (e.g. Kelley & Cranshaw, 2013; McKee, 2013; Rivers & Lewis, 2014). These guidelines are taken over to further prevent harm to the unknowing participants of this study. The suggestions from these studies are compounded into four overarching guidelines for ethical use of Twitter data; (1) creating transparency in methods of Twitter research, (2) anonymising the Twitter data, (3) not taking the data out of context, and (4) not tracking the Twitter users through different platforms. This study is keeping to these guidelines to minimize the potential damage to individuals and therefore, overwrites the need to inform the individuals of their participation and receive consent.

The first guideline is shown throughout the chapters by carefully explaining each step of the Twitter research process in this chapter and reiterating the methods used in each results chapter. This second consideration is applied to this study in various ways including making the sample members unidentifiable. For the protection of the identities of the people in this sample, this study only uses aggregates of people or anonymises them to avoid identity detection. This anonymization is seen throughout thesis even in the methods and appendixes. The third guideline of not taking tweets out of context relates to the initial reason for tweets being sent for the interest of their followers and not for the exploitation by third parties. Twitter has ‘user privacy expectations’ in which the user has consented with the terms and conditions but does not expect the data to be used for anything other than interacting on a social network site (Kelley & Cranshaw, 2013). Analysing tweets in isolation makes it difficult to keep the data in context as the context is unknown. Therefore, this third guideline is dealt with by taking the general Twitter activity into account and make conclusions based on the content of the tweet in connection with overall Twitter activity. For the final guideline on not tracking the participants through different platforms, this study relied solely on the information provided in the Twitter data and profiles. No attempt was made to contact the individuals, find them on other social network sites, or make them aware of their participation in this study.

This study was given a favourable opinion following a departmental level ethical review at the University of Portsmouth (a letter of approval and declaration of ethical conduct is included in Appendix A).
4.4 Creating the Twitter data for the present study

For the present study, it was necessary to find out which Twitter users have been in contact with The Filter Wales and the smoking-related tweets they have sent. This stage of data collation was undertaken with the assistance of the ASH Wales team and involved the compilation of two datasets; one set contains the descriptive information of the sample members, and the other contains their associated smoking-related tweets. The schematic depiction of the data collection and preparation process is illustrated in Figure 4.1. The remainder of this chapter explains the details of each part.

Figure 4.1 Diagram of the data collection process
4.4.1 Twitter as part of the Filter campaign

The Filter Wales uses Twitter to increase their ability to reach out to young people in Wales alongside other social media output and outreach events mentioned in the previous chapter (section 3.4). The Filter Wales’ social media team follows a series of discrete steps to find young people in Wales that tweet something about smoking. This is beside their daily tasks of presenting information in original tweets, retweets from other health organisations, and posts on the smoking-related news. The first phase of the Filter campaign is the gathering of tweets on the topic of smoking. They do this via search engines HootSuite and the Twitter Company. These search engines show all the tweets within the given parameters and with the selected words (i.e. smoking, tobacco, cigarettes, fag, baccy, rolly). For example, the team showed me how to set the search engine for ‘Cardiff +30 miles; cigs’ which indicates that the Twitter Company offered all the tweets with the words ‘cigs’ in it within a 30-mile radius of Cardiff. The Filter has been in contact with Twitter users within a geographical radius encompassing all of Wales. This method caused several places in England to be included as well, for example, North Somerset and Bristol which are so close to Newport and Cardiff that they have been picked up by the geographical parameters of the search engines. The exact geographical parameters were unknown as the social media team members did not remember when asked in 2017.

Subsequently, the team selects the tweets they find interesting and reply to, like, or retweet that specific tweet. As the Filter is an organisation designed to help young people (age 11-25), the social media team has to assess whether the Twitter user could fit in their target group. The age is guessed on the basis of the photo on their Twitter profile which has its limits as it could lead to some Twitter accounts being missed or wrongly added. The Twitter engagement of The Filter Wales is based on an ‘assumed’ target group, but that does not present particular problems as they help anybody who may require smoking cessation assistance. The social media team was also selective in their contact with users; not all smoking-related tweets were retweeted or replied to (specifically when they were offensive), and not all retweeted and replies were about wanting to quit smoking (other replies could concern negative experiences with others’ smoking).

4.4.2 Data harvesting of The Filter Wales Twitter

The creation of the two datasets started with a collection of all the Twitter activity via a request made to the Twitter company from the Filter team at ASH Wales. The Twitter Company allows for anyone to request the entire Twitter history of his or her Twitter account at any time at no costs. In the ‘Settings and Privacy’ tab on the Twitter profile,
there is a label with ‘Your Twitter data’. Once the Twitter user clicks on that and enters the password to their account, there is an option to download the entire Tweet history. Within a short time, the Twitter Company sends a folder with the Twitter data retrieved from an online database. This archive includes all the Twitter activity (i.e. tweets, retweets, and replies) of the account.

The Filter Wales requested this data folder and collated it for use in the research presented here. Figure 4.2 below is a screenshot of the data in CSV format opened in Excel and shows the elements which are associated with each tweet, i.e. the unique ID of the tweet, the timestamp, and extended URL with an attached picture or website link. The text in column F of Figure 4.2 contains the actual tweets.\footnote{the usernames in column F are not from ordinary individuals but organisations and health professionals.} This column was isolated from the dataset, and all the usernames (e.g. @Kickbuttsday) were copied out of that text into another file so that a list of only usernames was constructed. Afterwards, the duplicates were removed from the usernames’ list, as well as the usernames linked to organisations (e.g. @YMCACardiff) and celebrities (e.g. @GarethBale11). As outlined in previous chapters this study focuses on ‘ordinary’ young people and the removal of tweets linked to organisations or famous individuals retained that focus. The dataset contained 2703 unique individuals.
Figure 4.2 Screenshot of the raw Twitter data in Excel derived from the Twitter Company
4.4.3 The young peoples’ data harvesting

Once the list of all the young people’s usernames was created, the Twitter history of these unique users could be collected. The tweets can be accessed through a Twitter Application Programming Interface (API). An API is a guideline or interface which allows access to resources owned by another. For this study, an API program called ‘Mozdeh’ (Statistical Cybermetrics Research Group, 2014) was used, but there are countless other possible programs. The Twitter Company allows the collection of up to 3200 tweets from a single user via the API free of charge. However, more tweets could be composed of this group of young people contacted by The Filter Wales Twitter team by having several data collection points (July 2015, September 2015, January 2016 and June 20166). At each collection point, more usernames were added as The Filter Wales contacted new young people. The final set of Twitter data consisting of 4.8 million tweets (made between 13 November 2008 and 30 June 2016) was a compilation of the tweets collected at each point sorted by user.

Figure 4.3 is a screenshot of the data that comes from the Twitter API in a CSV format opened in Excel. As shown in Figure 4.3, the data consist of the username, ID, timestamp, and content of the tweet which is similar to the data file from the Twitter archive (Figure 4.2). The Twitter API program presents the added information of the location, source, original author, and the number of retweets. The username, timestamp, the content of the tweet, and location encapsulate the information needed for the analyses in this study.

6 The end of June 2016 was chosen as the last data collection point due to the end of The Filter Wales’ funding by the Big Lottery. Currently, The Filter Wales still operates due to other funds.
Figure 4.3. Screenshot of the raw Twitter data from the API program in Excel * The usernames (column A) and the first part of the tweet (column D) are altered to conceal the identity of the sample member and their Twitter network
4.4.4 Twitter profiles

Besides the Twitter archives of the individuals, all their Twitter profiles (2703) were collected on June 30, 2016, by making a screenshot of the profile of each username. These profiles provide additional information about the sample members such as their (presumed) gender, age, and Twitter activity.

Figure 4.4 illustrates a Twitter profile, and due to ethical consideration, it is my Twitter profile that serves as an example. The parts of the profile used in the analysis are on the left side of the screenshot underneath the photograph and also in the middle section which summarises the user’s Twitter activity. The information on the left consists of the username (@kim24501) which is unique, name (Cornelia van Diepen) which is in most cases the name of the user, and the date the Twitter profile was created (March 2015). The Twitter activity consists of how many tweets were posted by the user (50), how many people he or she follows (127), how many followers are following the user (60), how many other tweets the user has ‘liked’ (51), and lastly the relatively new ‘moments’ in which a user can select the posts that are particularly special to him or her (0).7

From information derived via the screenshots of the profiles, it became evident that not all of the sample members fitted within the definition of the target group for this research and some were eliminated at this point. For example, the personal message often revealed if the Twitter user was a health professional. Out of the original 2703 unique Twitter usernames described in section 4.4.2 above, 2186 (81.1%) were assumed to be young people in the Wales region. The 507 accounts that were excluded were organisations (e.g. schools, youth clubs), health or youth professionals, and celebrities. Another six Twitter users were excluded because their personal Twitter feeds consisted of less than 50 tweets. The Twitter API program could not always access the tweets because the Twitter users changed their privacy settings to a ‘protected profile’ or deleted all their tweets in the recent past. This deleting of tweets happened in 4 cases. If there were less than 50 tweets, a proper analysis would be difficult as tweets often do not contain much information on their own but need to be viewed in multitude to provide data detailed enough for research purposes (Bryman, 2012).

7 ‘Moments’ are not taken up in the results as this Twitter function did not exist when the profiles were collected.
Figure 4.4 Screenshot of a Twitter profile
4.4.5 Derived variables

Gender and age were essential variables for this study, and while they were not part of the profile, they could be deduced from it. Gender was derived from the name, personal message or profile picture. If these were inconclusive, a quick run through the tweets revealed the gender. Age was harder to deduce, and only 27.4% of the Twitter profiles showed a clear indication of the user’s age. The way to develop the age-variable consisted of the individual user putting their birthday on their profile or mentioning their age in the personal message. These ages were directly taken over in the age variable. At other times it could be learned through their username (e.g. @####93 which together with the profile picture could lead to a reasonable assumption that ‘93’ stood for 1993). After which the birthyears were transformed into the age the individual would be on the first of January 2016. These gender and age indications may not be completely accurate, and a limitation that must be acknowledged is that their estimation is based on a number of assumptions as outlined above.

All the tweets from the Twitter users over the different collection periods were gathered and uploaded into the ORACLE database server (ORACLE, 2011) to form one large dataset consisting of 4.8 million tweets from 2180 unique users. Moreover, the information about the Twitter profiles was gathered from these screenshots and catalogued in a table with all the usernames for this sample (see bottom right of Figure 4.1) which was created to provide information about the young people in the sample and their Twitter characteristics such as gender and Twitter activity. As data analysis proceeded, additional data items were added, and the details of these are described below. This data has been predominantly valuable for the examination of the sample population in Chapter 6 but were also relevant to all other results chapters as multiple parts of the analysis of the tweets incorporated gender and age differences.

4.5 Data preparation

The raw data from The Filter Wales Twitter feed and the sample’s Twitter feeds needed to be ‘cleaned’ before they could be used in the analysis. As mentioned previously, these millions of tweets consist of information beyond the content of the
tweet (such as Twitter user, timestamp, and geolocation). This next part is a description of how the data were prepared and which additional data were collected.

4.5.1 The smoking-related tweets dataset

For this study, the smoking tweets were extracted from the entire dataset of 4.8 million tweets (see diagram in Figure 4.1). The smoking-related tweet dataset was produced from the entire dataset that was uploaded into ORACLE (ORACLE, 2011) as this program can handle big data and is able to extract only the tweets which contained smoking-related terms. To gain only the smoking tweets from the 4.8 million collected tweets, SQL*Developer is a query interface tool that connects to the ORACLE database server. The terms that were used in the query interface were shaped by using the list of possible words the team at The Filter Wales used in the search for relevant tweets across the youth cohort, many other smoking-related words from smoking literature, and all other smoking-related words that were deemed relevant. The query is provided in Appendix B.

The selection of the definitive set of smoking-related words that made it into the final smoking-related table was based on their presence in the complete dataset; if there were ten tweets or more with that smoking reference, the term remained in the code to create the smoking-related subsample, if they did not, they were excluded. This minimum of ten tweets was chosen because a somewhat substantial number of tweets on a specific topic was necessary to assess the meaning of the smoking reference in the lives of young people. For instance, ‘blunt’ was a term included in the initial selection but only appeared six times and was therefore not used as a smoking reference. Essential for the final code was the acknowledgement that spelling errors are a common occurrence on Twitter. Therefore, the terms were designed to collect all spellings of the most common words, e.g. ‘cig-’ collected cig(s), cigarette(s) and cigar(s). Placing a ‘%’ before and after the search term make sure that every tweet with these letters in succession regardless of what characters came before or after are collected. This is convenient for Twitter research (e.g. in hashtags #shishalover). The final terms (see Figure 4.5) were put into a separate table. The ORACLE database server shaped a subset of smoking-related tweets that could be exported to a CSV file.
After the creation of the sample with smoking-related tweets, the next step was to go through them and delete all tweets without an actual reference to smoking. For example, ‘weed’ is not only a term denoting marijuana but also the past tense of ‘weeing’ which is not related to smoking. Similarly, ‘rolly’ is a term for a Roll Your Own cigarette but ‘prolly’ is a frequently used abbreviation for ‘probably’. These errors needed to be discovered in a manual check of all tweets, and once they were, the tweets not referring to smoking got omitted from the table. The sample resulted in a table of 16,688 unique smoking-related tweets made between 2 April 2009 and 30 June 2016 from 2180 individuals.

The entire dataset containing 4.8 million tweets was not checked for overlooked smoking-related tweets as this would be an extremely time-consuming activity without significant results as all the smoking-related terms have been tried in the previous step. Nonetheless, part of the analysis for Chapter 9 on the context of the smoking tweets concerns a subset of 50 randomly selected individuals to go through their whole Twitter feed, and no smoking-related tweets were found that were not identified before. The tweets written in the Welsh language were not considered as the researcher does not possess any Welsh language skills. The proportion of tweets in Welsh is unknown as the Twitter feeds do not give an indication of language. Consequently, the Twitter profiles were checked if any were in Welsh and none of them had a Welsh profile or tweeted in Welsh in their most recent posts.

4.5.2 Smoking status

The first variable that was extracted from the smoking-related table was the smoking status of the sample members. Smoking status was measured by taking the smoking-related tweets from each individual in the sample and reviewing the text content of their tweets in chronological order. The classification options are shown in Table 4.1. There is some caution to the labelling of smoking status as it is limited to the tweets that were available. Therefore, it is unclear if the persons’ smoking status on Twitter coincides with real life.
Table 4.1 Classification of the smoking status

<table>
<thead>
<tr>
<th>Smoking status</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-smoker</td>
<td>the sample member tweeted about smoking but not about smoking themselves</td>
</tr>
<tr>
<td>Smoker</td>
<td>the sample member tweeted about smoking themselves without tweeting about a serious quitting attempt</td>
</tr>
<tr>
<td>Quitter</td>
<td>the sample member tweeted about a serious quitting attempt and did not post tweets on relapsing or smoking in a later period</td>
</tr>
<tr>
<td>Relapser</td>
<td>the sample member tweeted about a serious quitting attempt but did post tweets on relapsing or smoking in a later period</td>
</tr>
</tbody>
</table>

4.5.3 Smoking type, person, and action

Young people engage in other forms of smoking products besides tobacco (i.e. marijuana, e-cigarettes, and shisha) and, as seen in section 2.5 of Chapter 2, these products are smoked for different reasons. Therefore, the smoking-related tweets in the table were categorised based on the kind of smoking product (listed in Table 4.2). All these terms were subject to the variance in spelling, and all tweets were reassessed for subtle expressions, e.g. “Lost my joking cigarette today now its hard” specifies an e-cigarette and not a tobacco one.

Table 4.2. Classification of the smoking product

<table>
<thead>
<tr>
<th>Smoking product</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobacco</td>
<td>tobacco, baccy, roly, smoking (without specifics), fag, Marlboro, cigarette &amp; cigar</td>
</tr>
<tr>
<td>Marijuana</td>
<td>stoned, marijuana, cannabis, weed &amp; splif</td>
</tr>
<tr>
<td>E-cigarettes</td>
<td>e-cig, fake fag, joking cigarette/fag &amp; vapour</td>
</tr>
<tr>
<td>Shisha</td>
<td>waterpipe, shisha &amp; hookah</td>
</tr>
</tbody>
</table>

The smoking product will affect how people are tweeting about smoking as they attach a different meaning to themselves smoking or seeing other people smoking in their proximity. A way for people to vent their frustration is through Twitter, and the social network sites are commonly used to express social disapproval (Kwak, Lee, Park, & Moon, 2010). Consequently, the tweets in the tables were categorised into who was performing the smoking activity, i.e. the person posting the tweet ‘First Person’ or anyone else ‘not First Person’. All tweets not following this pattern were
characterised as ‘other’. Examples of the category ‘other’ included tweets that had ‘smoking’ as the object or statements about new policies.

Lastly, these tweets were characterised by the type of action they contained as illustrated in Table 4.3. Different activities have meaning as to why they are posted at specific times. For example, more quitting smoking tweets are likely to be made in October as that is the time of the Stoptober campaign.

Table 4.3 Classification of the smoking activity

<table>
<thead>
<tr>
<th>Smoking activity</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking</td>
<td>The tweet mentions the action of smoking</td>
</tr>
<tr>
<td>Desiring to smoke</td>
<td>The tweet mentions the wish to smoke</td>
</tr>
<tr>
<td>Thinking of quitting</td>
<td>The tweet mentions wanting to quit/mentioning quitting in a later period</td>
</tr>
<tr>
<td>smoking smoking</td>
<td></td>
</tr>
<tr>
<td>Quitting smoking</td>
<td>The tweet mentions having quit smoking</td>
</tr>
<tr>
<td>Other</td>
<td>Any statement, retweet or tweet that was impossible to decipher in term of smoking activity</td>
</tr>
</tbody>
</table>

These elements are added to the smoking-related table and will feature in the results chapters. Dividing the tweets into broader themes as described above led to a more manageable overview of the table of 16,688 tweets and made it useful for the aim of this study namely to gain a better understanding of the meaning of smoking through tweet content. The analysis of this data is outlined in Chapter 5 section 5.2.2, and the output is applied to the results of (mainly) Chapter 7.

4.5.4 Geolocation

As mentioned previously, a feature of the tweets gained from the API program is the location from where the tweet is sent also known as geolocation. Geolocation refers to a technique in which geographical data about a person or device is derived from location information stored in digital format. The geocoded data is derived from devices that consist of a GPS (Global Positioning System) and is in most cases gathered from mobile devices such as mobile phones and tablets when the location application is switched on. This information is given in coordinates and indicates a place the sample members have posted tweets.

Geolocated tweets have a location specified as a pair of latitude and longitude coordinates attached to a single tweet. The grid reference provides a spatial
resolution with a 10 metres accuracy (M. Graham, Hale, & Gaffney, 2014). An example of the geolocation data from the API output is:

```
51.4691939,-3.1823309], "coordinates": {"type": "Point", "coordinates": [51.4691939,-3.1823309]}, "place": {"id": "4644af0995f5f5c7", "url": "https://api.twitter.com/1.1/geo/id/4644af0995f5f5c7.json", "place_type": "admin", "name": "Cardiff", "full_name": "Cardiff, Wales", "country_code": "GB", "country": "United Kingdom", "contained_within": [}
```

The coordinates of the geolocation signal straight at the location of the tweet originator (especially in relatively remote places). Most Twitter users are unaware of the private information they disclose, and with the addition of the geolocation, individuals can easily be identified. On average only 2-5% of all tweets are coupled with geolocation coordinates (Burton, Tanner, Giraud-carrier, West, & Barnes, 2012; Leetaru, Wang, Cao, Padmanabhan, & Shook, 2013). This number is low and when used in the analysis is likely to present skewed results (M. Graham et al., 2014).

In this study, 1484 tweets out of the smoking referenced table were geotagged and were posted by 462 sample members (21.2%), a proportion already higher than anticipated. The relatively high number of geolocated tweets is likely due to the use of geolocation in other applications predominately used by the young people. These might include applications such as Snapchat, Instagram, Tinder and Running apps that require the person to have their ‘location’ on their mobile phones switched on. If the ‘location’ is still switched on when they post something on Twitter, this geolocation information will be added to tweets.

To improve the chances of successfully determining the place of residence for even more sample members, all the (4.8 million) tweets collected from the entire sample population were used to increase the prospect of collecting multiple geolocation coordinates for the 2180 individuals. By applying the geolocation collection method from the entire dataset of 4.8 million tweets, the compilation of the geolocation increased to 394,775 geolocated tweets from 1698 Twitter users (84%) in this sample.

These geolocation coordinates could then be placed on maps through ArcGIS (ESRI, 2015). It is important to note that these geolocations do not necessarily indicate where the Twitter users live (as will be visible from the plotting of the maps in
Chapter 6) but merely indicate the places from which the sample members have sent tweets.

4.5.5 Determining place of residence

One of the aims of this study is to uncover the reach of The Filter Wales Twitter campaign. This aim includes a focus on the place of residence of the sample member which goes beyond the mere plotting of the geolocated tweets as they only illustrate where the tweets were sent from. For determining where the sample population lives, maps needed to be created with the location of the majority of geolocated tweets per individual. This section reveals how the determining took place.

For this study, the geolocated coordinates were placed in Lower Layer Output Areas (LSOAs) so that coordinates could be clustered within small areas. An LSOA relates to UK census geography and the reporting of official statistics. They are designed to be of a similar population size with an average size of 1500 people. There are a total of 1909 LSOAs across Wales (Office for National Statistics, 2014a).

All the available geolocation coordinates (394,775) were plotted on a map of Wales and England with a raster map of LSOAs. The geolocation and mapping for the place of residence of the young people were facilitated by a point in a polygon to make more sense in ArcGIS release 10.3.1 (ESRI, 2015). A digital shapefile of the LSOAs was made available from UK data service Census boundary data. The digital LSOA shapefile was overlaid with the geolocated points of the tweets, and the two layers of digital information were ‘joined’ so that each geolocated point had an LSOA code attached to it in an associated database file. All coordinates that fell outside of England and Wales were not accompanied by an LSOA and were removed from the table which resulted in a usable file of 379,315 tweets (96.1% of the total geolocated tweets) from 1695 sample members.

The output of the data is illustrated in the screenshot of an Excel-file in Figure 4.6. The number, username, latitude, and longitude were taken from the Twitter sample, and the LSOA code and name were attached to each row. The first column indicates whether the coordinates fall within an English or Welsh LSOA or if the coordinates fall outside of these countries (e.g. rows 5 to 7 and 10 to 13). To conclude, the geolocated points falling within an LSOA varies per sample member, and there are geolocations in both England and Wales.
Figure 4.6 Screenshot of the output of the ArcGIS join of the coordinates and LSOA layers in Excel.

Individuals may tweet from several different locations and, therefore, guidelines to determine the likely place of residence of the individual had to be devised. The LSOA found most frequently per individual user was selected as their residential location. A cut-off point of at least five tweets in this table was chosen to ensure that the geolocations were not accidental. A total of 1250 sample members had more than five geolocated tweets in the table and were assigned an LSOA which consisted of the majority of their geolocated tweets. Using LSOAs created the possibility to link characteristics of place to the sample population and moreover, produced aggregates of people.

Using geolocation in research is not completely reliable as the Twitter users might tweet a lot from a particular location such as school instead of home. Therefore, a validity check was done by taking only the geolocated tweets when they were sent
between September and May, on weekdays between 6 pm and 9 am, and all day during the weekend. These times were chosen as the person sending the tweets was most likely to be at home instead of at work, school, or their holiday location at these times. If the most prominent LSOA per individual was identical, this was considered valid. The validity check revealed a resemblance of 89.2% (the full table is attached in Appendix C).

After the LSOA had been determined, the sample members with a majority of geolocated tweets from England (199 individuals) were eliminated from the sample as further analysis into a place of residence, and the additional locational information was specific to Wales. This resulted in a group of 1051 sample members with a place of residence variable. The density and spread of the sample members across LSOAs were plotted using ArcGIS (ESRI, 2015) with close-ups of the areas with the most individuals from the sample which are presented in Chapter 6.

Also, the proportion of smoking-related tweets with a locational reference (through the person posting the tweets) were plotted on a scatterplot in SPSS (IBM Corporation, 2013) to assess the correlation between smoking-related tweets per percentage of the youth population in Welsh Local Authorities. This was done to examine if the likelihood of having been contacted by the Filter Wales is likely due to the high youth population in that area or because they tweet more about smoking.

Reach in the sense of this study’s aim exceeds the geographical place to connect to information on deprivation levels through the Indices of Welsh Multiple Deprivation (WIMD) 2014 (DCLG, 2015) and the rurality of the location through Rural-Urban Classification (RUC) 2011 (DEFRA, 2011). The WIMD is the Welsh government official measure for relative deprivation for small areas in Wales and is made up out of eight domains: income, employment, health, education, access to services, community safety, physical environment and housing (Welsh Government, 2014). Deprivation is the lack of access to opportunities and resources and the scores for each domain are combined to create the indices of multiple deprivation. These scores were split up into quintiles for the creation of levels of deprivation.

The RUC categorises a range of statistical and administrative components by settlement and related characteristics and consists of four urban (e.g. Urban city and town) and six rural categories (e.g. Rural town and fringe) (DEFRA, 2011). Each LSOA is attached to one of the ten categories and the RUC assists in examining the
variation of mainly social and economic characteristics. For a better overview of this variable, the classification was dichotomized into urban and rural in the last part of the analysis in Chapter 6.

### 4.5.6 Health risk behaviour concepts

Health risk behaviours and their connection to smoking have been expanded upon in section 2.6 of Chapter 2. The purpose here is to observe the other health behaviours within the smoking tweets content as examining health co-behaviours expands the understanding of the social meaning of smoking. The health-risk behaviour concepts utilised in this study were alcohol use, healthy eating, and physical exercise which are in combination with smoking often applied in large questionnaires such as the Welsh Health survey. The references to co-behaviours were found during the read-through of the tweets in the smoking-related tweets table and labelled as having another health behaviour present in the tweets. In order to increase the speed of this process, a vast array of behavioural health terms were ‘searched’ in the smoking table, and if they did not appear at all, the terms were eliminated. For the co-behavioural tweets, the following terms were applied which are demonstrated in the final ‘Classification of Health Behaviours table’ (Table 4.4).

**Table 4.4 Classification of Health Behaviours. Percentages (%) refers to the proportion of smoking-related table.**

<table>
<thead>
<tr>
<th>Health behaviour</th>
<th>Total tweets</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol</td>
<td>782 tweets</td>
<td>alcohol, vodka, beer, wine, cider, Strongbow, drink, drunk, sober, whisk(e)y, tequila, wasted, rum &amp; shot</td>
</tr>
<tr>
<td>Healthy eating</td>
<td>473 tweets</td>
<td>food, diet, weight, obes%, chubby, eat, fat, healthy, fizzy, choco%, stone, calorie &amp; lb</td>
</tr>
<tr>
<td>Physical exercise</td>
<td>195 tweets</td>
<td>gym, exercise, muscle, run, jog, yoga, workout, working out, cardio &amp; fit</td>
</tr>
<tr>
<td>Multiple health behaviours</td>
<td>105 tweets</td>
<td>Any tweet already classified by at least two other health behaviours</td>
</tr>
</tbody>
</table>

All tweets were checked to see if they referred to actual health behaviours. The health behaviours were marked in the smoking-related table, and this revealed that several tweets had a reference to more than one health behaviour in the tweet. The last row of Table 4.4 covers this group. There were false positives as tweets referred to, for example, ‘a fat spliff’ and ‘*I can finally fit a cigarette in my ear*’. Those were found as all the tweets were read again and checked if they actually referred to health
co-behaviour. If those were false positives, these tweets needed to be categorised as ‘not consisting of co-behaviour’ which was the majority of smoking tweets (16,057; 96.2%). The subsample of only smoking co-behavioural tweets consisted of 1340 rows of tweets.

4.6 Understanding the tweets

Before ending this chapter on the possibilities of Twitter and the data gathering and preparation process, this section is written to help the reader understand the tweets and the addition of non-verbal cues in the examples. It elaborates on the elements used in tweets and the way they are described in this thesis to understand their content.

Overall, the tweets are copied straight from the dataset; no words or grammar are changed. However, section 4.3 above expanded on the moral ethics to protect the Twitter users from harm that is caused by identification. Therefore, the usernames present in the tweets are anonymised by hashtags, i.e. ‘@#####’. Another alteration of the tweets is the deletion of profanities by inserting a ‘*’ in the place of a vowel. These insertions are all made by the researcher as none of the sample members self-censored their tweets in the same way.

4.6.1 Sarcasm and other sentiments

When reading the tweets, Twitter users often use sarcasm and jokes as they would in their face-to-face interactions. However, written words lack nonverbal cues that would assist in delivering that message. Consequently, the originator of the tweet needs to indicate in some way that the tweet is meant sarcastically or as a joke. Common features are the emojis mentioned in section 4.2.1. To indicate sarcasm or a joke, a ‘wink face ;) ’ is often applied:

“@##### I was about to accept your apology until I saw that wink at the end... below the belt. ;)

This example features an indication of joking as this tweet ends with a wink. Another option for revealing the tweet has humorous intent, is for people just to write it in the tweet:
“Dynamo the Magician is apparently in Wrexham .. some may say they hope he makes it disappear :p #banter #onlyjokes #wales”

“OMG guy just said to me "your makeup looks really good today" while i wasn't wearing makeup SWOON just kidding that didn't happen”

Other feelings are often revealed through phonetic expressions (e.g. “nomnomnom” or “bleh”). These phonetic expressions are valuable for understanding the meaning of the tweets from an audience perspective, and the sample population applied those expressions in significant quantities. In face-to-face conversations, these same sounds can also be used to aid the meaning such as “haha” or “naaaaah” and are an addition to the tweets so that the followers understand the intention:

“You actually fill up my newsfeed with your bullsh*t, you tweet every 5 minutes urgh”

“2 in the morning and the bf decides to bake cakes, what a weirdo grrrrrr I got work in the morning!!!”

As shown by the examples, the sarcastic intent or emotional load of the tweets submit to rules of using an emoji (mostly the wink), express how it should be read, or by adding in phonetics cues. If there is no indication of these signals, the intention of the tweet should be understood by the word content as there is no evidence to the contrary. The tweets without a ‘sarcastic’ reference were analysed as ‘serious’.

4.6.2 Understanding retweets

A number of tweets in the smoking dataset are retweets (1235, 7.4%). These retweets are difficult to interpret for analysis because they are not created by the individual in the sample. Most retweets inform the followers of the retweeter about an interesting communication made by someone else (e.g. a news article), but part of the retweets are also made to transfer the attention onto the retweeters (e.g. a witty pun or joke) (Boyd, Golder, & Lotan, 2010).

Another form of retweeting is done by the sample members in which ‘original’ tweets were copied from others without the acknowledgement (retweet element) of the originator and were discovered solely because they were identical. Often those texts were taken from the same popular source, generally a pop culture reference. For

---

8 The underscore was inserted by the researcher to emphasis the important element in the examples.
example; “Lighting a cigarette and wishing the world away” was used as an original tweet by four people in the sample. This tweet is a lyric from the song called ‘Two fingers’ by singer Jake Bugg.

Therefore, for the interpretation of the retweets and ‘non-original’ tweets, those in which the original tweeter posts an opinion or experience are interpreted as if the users copied it in an attempt to express their feelings through someone else’s words. However, when the retweet is a statement, tweets from a celebrity, or quotes from entertainments shows, they are not valuable for understanding the social meaning of smoking, and they are classified as ‘other’.

4.7 Concluding remarks

This chapter has outlined what Twitter is, how it is used in health research, and described how using Twitter as a research tool differs from traditional methods. It further emphasised that there are no established ethical limits just guidelines to using Twitter as a research tool, but for this study, measures are taken to prevent identifying the individual who created a tweet. The second half of the chapter reported on the data gathering and preparation of the data for the analysis, and through the use of geolocation, the chapter outlined how the original dataset can be enhanced by linking it to data about ‘place’ context. Lastly, this chapter included a guideline on how to understand the intention in Twitter content. This next chapter reiterates the aims and objectives of the study and formulates the traditional and novel methods of analysis.
Chapter 5. Methods of Analysis

This chapter builds on the introduction to Twitter research and the discussion of data assembly for this thesis that were set out in the previous chapter and is an account of the methods of analysis used to address each of the objectives by demonstrating their practice and place in this Twitter-based study. The methods described in this chapter cover a range of approaches including quantitative data summary and qualitative analyses of the text contained within tweets. Attention also focuses on more innovative techniques designed for the analysis of sentiment and general use of linguistic markers within social media information. Traditional and Twitter-specific methods are elaborated on in this chapter to create a full picture of the potential of Twitter for exploring the reach of The Filter Wales and the social meaning of smoking.

The chapter begins by outlining the descriptive methods which include tabular representations of demographic and smoking status and computer-generated maps of the geolocated information (5.1). This section further explores patterns by applying quantitative content analysis, temporal analysis, and sentiment analysis to the Twitter data. The second part of the chapter (5.2) outlines the qualitative methods applied to the data in the form of qualitative content analysis, analysis of the discourse, and linguistic analysis.

5.1 Descriptive methods

This first section of the methods of analysis outlines the descriptive methods used for the objectives and starts with the tabular representations. Descriptive tabular analyses were used to investigate the gender and age differences in combination with each other and other variables, e.g. Twitter activity and smoking status (all of which are described in section 4.5 in Chapter 4). These frequency distributions are purely to tally counts and where appropriate the patterns in these contingency tables are tested using chi-square ($\alpha=0.05$). This was used to address the first objective of this thesis;

1. to identify the gender and age differences in the use of Twitter across this study sample.
5.1.1 Locational depictions

Another purpose of the descriptive analysis is to examine the place of residence of the young people in the sample by achieving the second objective;

2. to evaluate the reach of the Filter Twitter campaign through an exploration of the geolocation of the Twitter users.

This objective required a straightforward presentation of Twitter engagement according to the geolocation of the tweet. The locational information (detailed in sections 4.5.4 in Chapter 4) was used to plot the geolocated tweets on digital maps using ArcGIS (ESRI, 2015), i.e. on a world map, a map of the UK, and a map of Wales. This showed the global extent of geolocated tweets of the sample and the patterns of tweeting across the UK and more specifically, Wales.

Furthermore, the place of residence of the sample members was determined and displayed on a map of Wales in which the distribution of the sample population is placed within LSOAs. Comparing the residential location maps for Wales with the raw Twitter activity maps showed how, even though people tweet from many locations, the place of residence is directed to a few geographical regions from where the individuals in the sample tweet the most.

To further analyse the reach of The Filter Wales, a scatterplot was created to plot the number of smoking-related tweets according to the 22 number of local authorities within Wales against the total youth population (aged between 11 and 25) per local authority. This plot was undertaken to explore whether smoking-related tweeting activity was over- or under-represented in any local authorities. The association was assessed using Pearson’s correlation coefficient (α = 0.05).

Individual LSOA places of residence (i.e. most common LSOAs per sample member) were connected to the Welsh Indices of Multiple Deprivation (WIMD) (Welsh Government, 2014) and a Rural-Urban Classification (RUC) (DEFRA, 2011) for measuring deprivation level and rurality for individual sample members, respectively (see section 4.5.5 of the previous chapter for details on the WIMD and RUC). The results are presented in tabular form to show the distribution of the sample members across these two dimensions of place. This descriptive data illustrate the socio-economic differences in the sample population and the reach of
The Filter Wales. It also provided valuable information about reach in more rural locations across Wales.

Further descriptive tabular analyses were used to investigate the distribution of men and women and smoking status among quintiles of deprivation. The same was done for the rural/urban classification, and where applicable, the chi-square test ($\alpha=0.05$) was used to examine the statistical significance of any associations that are revealed within the tables.

5.1.2 Quantitative content analysis

Moving on from the report on the sample population, a quantitative content analysis was done with a description of how often specific themes in tweets appeared. “Content analysis is an approach to analyse documents and texts that seek to quantify content in a systematic and replicable manner” (Bryman, 2012, p.289). It was deemed necessary to take an automated quantitative content approach. Reading through 16,688 rows of data and assigning codes manually would have been too time-consuming and the sheer volume of data may have threatened the reliability of this part of the research. Therefore, part of the analysis follows a method where the number of references to predefined codes is counted using Excel’s select application. A quantitative content approach is the most common methodology among social research concerning secondary data (Bryman, 2012) and has also been widely used in Twitter-based research, e.g. Carvazos-Regh et al. (2015) researching marijuana tweets by counting the references to the person, the action they are performing and if it contains a reference to pop-culture or legislation. Another example would be Harris et al. (2012) on tweeting about health regulations for e-cigarettes as they counted the number of tweets per preselected topic (e.g. safety, lies, science, flavour) and analysed manually if the tweet was positive or negative. Both studies are mentioned in more detail elsewhere in this thesis. The quantitative content analysis addresses the following objective;

3. to examine the engagement with the Twitter feed, including differences in the content of tweets posted.

The codes have been defined in Chapter 4 section 4.5.3 and relate to the object of the text (tobacco, marijuana, e-cigarettes and shisha), the subject of the text (First Person or not First Person), and the activity (smoking, desiring to smoke, thinking of
These predefined themes present the most general content analysis by defining the object, subject and verb, and reveal the importance of particular topics in the posts of the sample members.

5.1.3 Temporal analysis

The coded data from the quantitative content analysis is further scrutinised through a time lens. The moment a tweet is posted gives a significant indication of the social meaning young people give to smoking via Twitter and it presents an opportunity to place the content in a timeframe (Dodds, Harris, Kloumann, Bliss, & Danforth, 2011; Golder & Macy, 2011; J. H. West et al., 2012). It reveals not only what opinions the sample members express about smoking but the time analysis reveals when they find it necessary to post it online. Moreover, instead of asking people about what they think of smoking, retrospectively, the timestamp (e.g. Mon Oct 21 17:19:55 +0000 2013) on the tweets reports the exact time and date when the thought was posted.

The statistical program SPSS (IBM Corporation, 2013) was used in the temporal analysis for plotting the number of tweets made between 1 July 2013 and 30 June 2016 on time series graphs, with hours of the day, days of the week, and months of the year. This part of the analyses addresses part of objective number four:

4. to examine the temporal and sentimental qualities of the smoking-related tweets.

An important part of the temporal investigation is to compare the timing of general Twitter activity alongside the timing of smoking-related tweets. It is possible that smoking tweets could be composed and posted during the individual’s regular Twitter activity and would perhaps indicate that the individual does not attach special meaning to these smoking-related tweets. However, if a specific temporal pattern is revealed across smoking-related posts, then this could provide additional insights into the specific time of the day that individuals might think about smoking and reveal time-related triggers that lead to such activities (e.g. expressing discontent over new tobacco control regulations). Similar research was done by West et al. (2012) who studied problem drinking references on Twitter. They reported that the number of tweets about alcohol increased during real-time events such as Halloween and New Years’ Eve. This study revealed how the smoking Twitter activity is affected by events and patterns in the outside world by examining where smoking tweets differ from the general Twitter activity. Once those places were identified, the
smoking tweets can be related to real life events that could have caused that increase. Therefore, a selection of 590,296 ‘general’ tweets by 233 sample members was selected to illustrate the overall Twitter activity. Comparing one eighth of the entire dataset of 4.8 million tweets allowed for a representative number of tweets to portray the general activity. The sample members were randomly selected to achieve 12.5% of the entire dataset of 4.8 million tweets while keeping all the gathered tweets of the Twitter archives of these selected sample members. This puzzle of reaching 12.5% of the tweets by compiling the gathered tweets of selected sample members lead to this number of 590,296 ‘general’ tweets. The timelines of the smoking-related tweets were compared to those of the general Twitter activity to reveal possible patterns and variations.

Finally, the temporal qualities of the smoking tweets are analysed by plotting all the tobacco-related tweets on a timeline for the weeks in a year taken from data made between July 2013 and June 2016 to show a more cohesive representation of tobacco-related tweets over a longer period of time and to illustrate particular relevant dates in a year (e.g. the first of October indicating the start of the Stoptober campaign).

5.1.4 Sentiment Analysis

The type of content analysis described previously does not disclose any sentimental meaning within the tweets. The analysis described merely documents the prevalence of tweets across different content categories. A machine learning program can add some weight to the meaning of the content of the tweet by automatically classifying tweets according to their direction and strength of sentiment and as such address the sentiment dimension of objective four listed above. Machine learning is a field of computer science in which computer programs take over tasks and is convenient to use for sentiment analysis on a large quantity of data as it improves the specific performance in a fast and unbiased manner.

Sentiment analysis is a linguistic process in which the text in the tweet is tested on words with negative or positive associations by taking into account all sentiment markers such as emojis and swear words. The advantage of using sentiment analysis is that the general intent of the tweet content can be illustrated. Therefore, many studies have applied sentiment analysis to their Twitter-based data which they have gained through a variety of options, e.g. manually (Myslín et al., 2013), R-packages
(Ghosh & Guha, 2013), client-side JavaScript (Signorini et al., 2011) or crowdsourcing (Cavazos-Rehg et al., 2015).

The creators of the API Mozdeh (Statistical Cybermetrics Research Group, 2014) have also created a free sentiment analysis program called ‘Sentistrength’ (sentistrength.wlv.ac.uk). Sentistrength is open source software that can be downloaded from their website. The program provides a systematic way of extracting positive and negative intent from pieces of text as they can appear together within the same tweet (Thelwall, Buckley, & Paltoglou, 2012) which is uncommon for sentiment analysis programs. Moreover, this machine learning program involves no intra- or intercoder bias (in contrast to coding manually) as there are no people with personal sentiments involved and it does not require any programming knowledge (in contrast to R-packaging or Java script). Crowdsourcing was considered but was dismissed due to ethical concerns that random people could access the tweets out of context.

Sentistrength is used to measure the sentiment within each tweet, and it has been widely applied to measure sentiment across this social media data (Bhattacharya, Srinivasan, & Polgreen, 2014; Thelwall et al., 2012; Thelwall, Buckley, Paltoglou, & Cai, 2010). However, a limitation of the program is that it only gives a score towards positive and negative sentiment, but not to the kind of emotion. Moreover, the sentiment analysis is sensitive to (i.e. unable to read) incorrect grammar and ‘fashionable’ abbreviations (for example, people writing ‘obvs’ as an abbreviation of ‘obviously’). This is a critical challenge in Twitter feeds from young people as, especially with the character limit of Twitter, they are more likely to apply their own abbreviation that are not used (yet) in everyday conversations. Another critical challenge relates to the meaning young people give to certain words which may differ from the official definitions such as the word ‘sick’ meaning ‘cool’ rather than ill. Applying sentiment analysis manually would prevent these limitations but would also be prone to intra- or intercoder bias. These limitations could also be avoided by creating the machine learning code in R, JavaScript or Python, but those skills go beyond the capabilities of the researcher. Therefore, the limitations of the use of Sentistrength in this study are taken into account for the results and the discussion.

The way it works is that for every tweet, two scores are given: positive and negative. The minimum score given to each piece of text is (1, -1) and with each word with a
positive or negative connotation, a point gets added to the score. The maximum sentiment score for each tweet is +5 and -5 indicating the content is extremely positive and extremely negative. These apparently polarised terms do coincide and can occur in parallel (Thelwall et al., 2012) as the following example from this study illustrates:

"@###### someone's stolen my tobacco and I'm f*cking pissed off as f*ck. I need to love you pls x" (3, -5)

This is broken down by Sentistrength:

@###### [0] someone’s [0] stolen [-1] my [0] tobacco [0] and [0] I’m [0] f*cking [0] pissed [-2] [-2 LastWordBoosterStrength] off [0] as [0] f*ck [-2] [[Sentence=5,1=word max, 1-5]] I [0] need [0] to [0] love [2] you [0] pls [0] x [1]

This example of the sentiment analysis scored a positive score of 3 because of the words ‘love’ and ‘x’ and scored a negative score of 5 because of the words ‘stolen’, ‘(f*cking) pissed’, and ‘f*ck’. This method of analysis was done by using the Sentistrength program for each of the tweets in the smoking referenced table. By itself, the sentiment scores of single tweets do not mean much, but average scores of a large group of tweets on the same topic will show an overall sentiment towards that subject (Thelwall et al., 2010). The average scores are applied to the outputs of the quantitative and qualitative content analysis.

The sentiment scores are attached to each tweet which makes it possible to use the temporal information and plot the sentiment score on timelines (i.e. hours, weekdays, and months). This approach is similar to Golder & Macy (2011) who researched at which times of the day positive or negative sentiment appeared most frequently. They looked at tweets from 14.3 million Twitter users around the world and found that overall weekends and periods with longer daylight had the most positive sentiment while during the night hours, negative sentiment was more common (Golder & Macy, 2011). This part of the analysis is done to see if these results are similar for smoking-related tweets. The average sentiment scores for each type of smoking product was plotted on the timelines combining sentiment with temporal analysis. This analysis illustrates when the sample population is most positive and
negative about smoking to better understand the circumstances in which different meaning is given to smoking.

5.2 Qualitative methods

In addition to the graphical and tabular descriptive analyses, and the quantitative approaches to content, time qualities and sentiment, further in-depth qualitative analyses were applied to the Twitter data. These methods take on a qualitative approach to the Twitter data by providing an in-depth view of the use of Twitter by the sample members and illustrating the content of the tweets. Even though the data consist of many microblogs, there is a possibility to do qualitative research on the tweets when they are based on a theme or analysed as a whole. Qualitative analysis can be done on tweets and is necessary to gain a better understanding of the social meaning of smoking. The traditional qualitative methods described here include qualitative content analysis and the analysis of discourse. A novel method of linguistic analysis was also applied to the data and described at the end of this section.

5.2.1 Qualitative content analysis

The first qualitative method is another version of content analysis. Qualitative content analysis is an approach that emphasises the construction of the meaning of and in texts. Moreover, this approach allows the categories of the content analysis to emerge from the data and recognises the significance of the meaning given by the sample members (Bryman, 2012). The technique consists of identifying specific characteristics in texts. Most studies on content analysis on Twitter have a ‘codebook’ in which the researcher first searches for the codes in the relevant Twitter data and afterwards further categorises within the sample (Marwick, 2013).

The empirical studies mentioned in section 4.2.1 of Chapter 4 have applied a similar approach to the one attempted here. For example, Heavilin et al. (2011) coded each tweet with a dental pain reference until they reached thematic saturation and Sullivan et al. (2012) in their study on concussions gave each tweet a more general theme, e.g. ‘news’, ‘advertising’, and ‘downplay’ after which they analysed the tweets per code to uncover their deeper meaning.

It is difficult to achieve a comprehensive content representation in a set consisting of 16,688 microblogs. Therefore, a subtheme of health co-behaviours was chosen which
had two purposes. Firstly, a focus on health co-behaviours within the smoking-related tweets reveals the place smoking has in reference to other health behaviours for young people. Secondly, using a subtheme enables the researcher to probe more in-depth into the content of the tweet. Smoking is not an isolated behaviour, and by studying health co-behaviours, the meaningful connections between smoking and alcohol use, physical activity, and healthy diet become clearer. Overall, this qualitative content analysis created a more in-depth picture of the social meaning of smoking. The fifth objective is related to the content analysis of the smoking co-behavioural tweets:

5. to evaluate the extent to which co-behaviours are present in the smoking-related tweets.

To achieve this subsample of health co-behavioural tweets, each tweet in the smoking-related tweet sample was read and flagged if the tweet contained another health behaviour. There was a total of 1340 tweets that were flagged and included other health behaviours within the same tweet. The content analysis of co-behaviours was a process of interpreting the content and the underlying meaning of how these behaviours are presented and interplay within the tweets. This is an inductive process in which each co-behavioural tweet was coded with reference to the content (e.g. ‘buying alcohol and cigarettes’ or ‘losing weight through smoking’). The coding of the co-behavioural tweets led to a number of themes; 11 for alcohol co-behaviour, 11 for healthy eating co-behaviour, 6 for physical exercise co-behaviour, and 5 for multiple co-behaviours. All of these codes are presented in the tables of Appendix D. For the final results, these themes were reduced to three themes per health behaviour. This was the moment when it became clear that healthy eating co-behaviour and physical exercise co-behaviour had the same overarching three themes and were combined in the results. The content analysis was further applied to the tweets with multiple health co-behaviours and having them previously coded these could be seen as a whole and presented in their own section in the results.

Lastly, to aid the in-depth interpretation of the health co-behaviours, a sentiment analysis (explained above in section 5.1.4) was applied to the subsample tweets to exemplify with what sentiment the sample members posted about the different health co-behaviours.
5.2.2 Analysis of the discourse

Another traditional qualitative method that was applied to the data was the analysis of discourse. Discourse analysis is an approach to find the meaning of a text (Coulthard & Coulthard, 2014; Schiffrin, Tannen, & Hamilton, 2005). This means uncovering not only what is being written but specifically ‘how’ it is written (Schiffrin et al., 2005). Discourse analysis is a specific methodological tool aimed at uncovering the underlying and unspoken power relations buried in texts. However, the Twitter data does not allow for such extensive examination. Therefore, this study applies an analysis of the discourse instead which aims at understanding the tweet content as part of an online discourse. The analysis of discourse will realise the following objective;

6. to identify the wider context of smoking for young people as evidenced by their Twitter archive.

For the analysis of discourse, the other tweets of the sample members needed to be examined to uncover the specific intentions of the smoking-related tweets and how they fit within the larger Twitter activity. Therefore, the analysis followed a systematic approach with a sample of fifty randomly selected young people out of the complete sample population (SPSS has a random selection function for this purpose). A subsample of 50 was chosen for an in-depth analysis of the young people’s Twitter use, paying specific attention to the patterns that emerged from an examination of the Twitter history. This created the possibility to build a narrative of these fifty selected young people.

There are 113,569 tweets gathered from the subsample of 50 and the analysis proceeded by focusing on the smoking-related tweets in the Twitter history of these subsample members. The smoking tweets (342 in total) were identified within the Twitter feeds and the days surrounding the smoking tweet were highlighted. The data was coded by the themes that would appear in the texts and how the participant wrote about those themes through the qualitative analysis program Atlas.ti 7 (Scientific Software Development, 2013). Atlas.ti is a program that assists in the qualitative analysis of large bodies of textual, graphical, audio and video data and offers a variety of tools to systematically code and provide an overview of the relevant content from each Twitter archive for each code (Scientific Software Development,
To give an example from the studies referenced in this research, Akre et al. (2009) applied Atlas.ti to their qualitative analysis of marijuana and tobacco co-use. Atlas.ti allowed the codes to be grouped in different hierarchical levels to be classified and formed four overarching themes to structure the results (Akre et al., 2009). The meaning and patterns within the Twitter archives of this subsample of 50 could be revealed by utilizing this program.

5.2.3 Linguistic analysis

Next to the analysis of discourse, a linguistics program was used to uncover the emotions and online persona of the individuals in the subsample of 50 and examine how their smoking tweets fit within the general Twitter activity. For example, upbeat smoking tweets might relate to being excited about smoking or tweeting upbeat posts in general. This linguistic analysis on Twitter data is a more complicated approach than sentiment analysis and focuses on individual users’ Twitter history which made the process more extensive and was, therefore, only applied to the subsample from the analysis of discourse.

A LIWC application called AnalyzeWords (Booth, Pennebaker, Pennebaker, & Wilson, n.d.), a free online program, can analyze a collection of tweets available from a single Twitter username to indicate the online personality. The words people use in everyday life are full of personality and emotion, and this program can examine these words and emojis quickly and efficiently to uncover particular styles and discourses within the Twitter feeds. AnalyzeWords is designed as a cooperation between two scientists and two reporters interested in Twitter language to provide the public with a free tool to uncover peoples’ Twitter style. The Twitter style is determined by three main subjects; Emotional style, Social style, and Thinking style. Table 5.1 illustrates the 11 subdivisions of the styles and the description given by the program.

---

9 LIWC application are generally expensive
Table 5.1 Categories of the linguistics analysis described by the Analyzewords program (Booth et al., n.d.)

<table>
<thead>
<tr>
<th>Category</th>
<th>Description from Analyzewords.com</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emotional Style</strong></td>
<td></td>
</tr>
<tr>
<td>Upbeat</td>
<td>Lots of positive words and ‘we’ talk. You’ve got energy, kid, if you ranked high on ‘upbeat’</td>
</tr>
<tr>
<td>Worried</td>
<td>Anxious language dominates your tweets as do nervous questions</td>
</tr>
<tr>
<td>Angry</td>
<td>Short of constantly writing in all caps, someone high in the ‘Angry’ category uses hostile words and talks a lot about YOU</td>
</tr>
<tr>
<td>Depressed</td>
<td>Sad, melancholy, inward-looking. Lots of self-references and frequent use of depressive words</td>
</tr>
<tr>
<td><strong>Social Style</strong></td>
<td></td>
</tr>
<tr>
<td>Plugged in</td>
<td>Socially engaged. A category reserved for prolific Tweeters; you scored high in this area if you use social words (’party!’) and include frequent shoutouts to your @friends</td>
</tr>
<tr>
<td>Personable</td>
<td>Engaged in other people’s well-being and at peace with expressing your own uncertainty about the world. High scorers in ‘personable’ use positive emotion words, ask questions, express their own ambivalence and reference others frequently</td>
</tr>
<tr>
<td>Arrogant/Distant</td>
<td>Well-read and smart with an arms-length approach to socializing. You scored high in this category if you discuss actions instead of emotions, use big words and don’t reference yourself much</td>
</tr>
<tr>
<td>Spacy/Valley girl</td>
<td>Excitable! If you rank high in valley girl; you love recounting your newest story with a lot of LOLLLLLLs!!!!!</td>
</tr>
<tr>
<td><strong>Thinking Style</strong></td>
<td></td>
</tr>
<tr>
<td>Analytic</td>
<td>If law school exams were a person, they would rank real high in this category. Ample large words and phrases that include complex thinking (e.g. ‘if – but not.....’)</td>
</tr>
<tr>
<td>Sensory</td>
<td>A tendency to reference your feelings and surrounding environment. A ‘Northern California’ approach to tweeting (sans reusable bag)</td>
</tr>
<tr>
<td>In-the-moment</td>
<td>Grounded in what’s hot now, with tweets that breezily reference today</td>
</tr>
</tbody>
</table>

To exemplify the output of this program, Figure 5.1 presents my online personality on Twitter through the AnalyzeWords LIWC. I have a high score on ‘Upbeat’ and ‘Worried’ meaning that I use positive words in my tweets and convey emotion. On the contrary, I have a low score on ‘Analytic’ and ‘In-the-moment’ which suggests that I do not use much punctuation in my tweets nor produce many hashtags. My online personality is positive but not very specific to the events of the day which is accurate as I use Twitter for academic purposes.
The linguistic analysis is presented with a division in smoking status (i.e. smoker, non-smoker, quitter, and relapser) to assess the variation in Twitter Style. Moreover, this LIWC is later on applied to the results of the analysis of discourse to uncover the association between Twitter Style, smoking status and the online context and discourse of the smoking tweets. This linguistic approach adds perspective to the understanding of these smoking tweets.

![Analysis of tweets from kim24501](image)

**Figure 5.1. Screenshot of the output of the AnalyzeWords program**

### 5.3 Concluding remarks

This chapter outlined the objectives in combination with the methods of analyses for the subsequent results chapters. This study applies approaches from both novel and traditional methods, outlining a wide range of possibilities for Twitter research. These methods together reveal the reach of The Filter Wales campaign and a better understanding of the social meaning of smoking from a novel angle while maintaining some classic elements of social research.

To begin the results section of this study, the next chapter describes the reach of The Filter Wales and characteristics of the sample population regarding gender, age, smoking status, and place of residence to provide a base for further research into the content and context of the smoking-related tweets.
Chapter 6. Describing the study sample

This first results chapter aims to describe the demographic, assumed smoking status and geographical characteristics of this sample to obtain a more detailed picture of the adolescents and young adults engaging with the Filter Wales campaign on Twitter. Therefore, this chapter examines the personal information gathered from the Twitter profile, the smoking-related tweets from each individual, and the geolocation attached to tweets. The specific research objectives that are being addressed in this chapter are:

- to identify the gender and age differences in the use of Twitter across this study sample
- to evaluate the reach of the Filter Twitter campaign through an exploration of the geolocation of the Twitter users.

The profiles of the 2180 young people described here have all either been retweeted by or have received a reply from The Filter Wales on one or more of their Twitter posts. This is the sample of individuals who are most likely young individuals from Wales. The evaluation by Meek, Hurt & Grant (2015) of The Filter Wales program did not consider, in depth, the demographics of the group or their smoking status and neither did they investigate the geographical reach of the campaign. Therefore, this chapter provides greater insight into the profile of these young individuals and demonstrates the personal information that can be gained from a deeper scrutiny of their Twitter archives and profiles.

In this chapter, the profiles of the Twitter users are used in combination (where possible) with the geolocation that is attached to the tweets. Section 6.1 starts with a description of the data derived from the Twitter profiles of the study group. This first section includes a description of the gender make-up and the age profile of the study group in combination with smoking status. The second section (6.2) focuses on the spatial distribution and socio-spatial geography of the Twitter users. Maps are produced which show the locations from where tweets are posted and displayed at a global level, a UK level, and finally, the detail across Wales is examined. The geolocation information is subsequently used to attach Lower Super Output Areas (LSOAs) as their probable LSOA of residence. Section 6.2 also discusses the distribution of smoking-related tweets of the sample members and the youth
population on a Lower Authority (LA) level to examine the reach of the Filter to areas with high youth populations. In the third part of the chapter (section 6.3), the LSOA place of residence is used to link the Twitter information to other geographical information, i.e. 2014 Welsh Index of Multiple Deprivation (Welsh Government, 2014) and the 2011 Rural/Urban Classification (DEFRA, 2011). This will reveal two characteristics of the place of residence which help contextualise the study sample and associated smoking behaviour.

6.1 The gender, age and smoking status of the sample

A summary of the information derived from the sample members’ online Twitter profiles is recorded in this section. Out of 2180 individuals, there are 1234 (56.6%) female and 946 (43.4%) male sample members. It was possible to gain the age of 597 (27.4%) of the sample population (see Chapter 4 section 4.4.1 for more details on how). The gender proportions are not similar to the proportion of young people found in Wales, wherein 2015, 50.8% of young people (i.e. aged between 11 and 25) were women and 49.2% men. Therefore, the following results are only applicable to this specific study and in relation to the individuals who came into contact with The Filter Wales’ Twitter account.

6.1.1 Age and gender

Table 6.1 reports the age and gender summary characteristics derived from the Twitter profiles. The age distribution is not normal, and so the median age was chosen as a measure of central tendency instead of the mean. As Table 6.1 indicates, the median ages and age ranges across both genders are similar, with the median age for men being just one year older than that for women and the age ranges differing by only a couple of years as the men have range of 20 years and the women’s ages range over 16 years.

Table 6.1 Summary data on age and gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number of people (%)</th>
<th>Median Age</th>
<th>Minimum Age</th>
<th>Maximum Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>273 (45.7)</td>
<td>22</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td>Women</td>
<td>324 (54.3)</td>
<td>21</td>
<td>15</td>
<td>33</td>
</tr>
</tbody>
</table>
The histogram below (Figure 6.1) provides additional detail on the age and gender makeup of the sample. This figure gives a more explicit indication that the distribution of the women’s ages tends to be towards the younger end of the scale compared to that of the men; the proportion of women 20 years or younger is 42.9% while this is only 31.1% for men. Moreover, the number of young people engaging with the Filter Twitter feed over the age of 25 is small for both genders. The bulk of the group's age lies between 17 and 23 which fits with the Filter’s age focus group (11 to 25-year-olds).

![Histogram showing age distribution by gender](image)

**Figure 6.1** The age distribution of the study sample according to gender.

### 6.1.2 Twitter activity according to gender

Table 6.2 summarises the Twitter engagement of the sample regarding the median number and range of tweets posted across the individuals. The table also indicates the median and range values for how many other Twitter users they follow (i.e. followings), how many followers they have (followers), and the number of ‘likes’ they have given to the tweets of others (likes). All this data is derived from the screenshots of the 2180 sample members’ Twitter profiles collected on 30 June 2016.

They are not based on the dataset of gathered tweets from these sample members as these only contained approximately 3200 tweets gathered retrospectively from 30 June 2016. It is clear from the Twitter engagement in table 6.2, that the sample population has tweeted a lot more than the 4.8 million of tweets API program
collected. The information in this section provides some sense of the range of Twitter activity and engagement across the sample. The table, for example, indicates that there is at least one sample member who has only tweeted 8 times\textsuperscript{10} and at least one member who has tweeted on 275,000 occasions. Reaching a quarter of a million tweets is possible when the individual starts conversations as each reply is considered a tweet.

Table 6.2 Twitter engagement across the sample

<table>
<thead>
<tr>
<th>Indicator of Twitter engagement (persons in sample)</th>
<th>Summary statistics for engagement activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
</tr>
<tr>
<td>Tweets (2180)</td>
<td>5874</td>
</tr>
<tr>
<td>Following (2179)</td>
<td>425</td>
</tr>
<tr>
<td>Followers (2180)</td>
<td>1137</td>
</tr>
<tr>
<td>Likes (2165)</td>
<td>947</td>
</tr>
</tbody>
</table>

There is a significant disparity between engagement of the individuals with Twitter as the variance between the highest and the lowest number across all engagement categories is quite large. Tables 6.3 to 6.6 summarise the total number of tweets users post (Table 6.3), the number of Twitter followers any one individual has (Table 6.4), the number of people a person is following on Twitter (Table 6.5) and the number of ‘likes’ the individual gives to other posts on Twitter (Table 6.6). All of these are broken down according to gender.

Table 6.3 Summary data on the number of tweets made by the sample

<table>
<thead>
<tr>
<th>Gender (persons in sample)</th>
<th>Total number of tweets</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (946)</td>
<td>9445963</td>
<td>5458</td>
<td>37</td>
<td>236000</td>
</tr>
<tr>
<td>Women (1234)</td>
<td>13496152</td>
<td>6272</td>
<td>8</td>
<td>275000</td>
</tr>
</tbody>
</table>

Table 6.4 Summary data on the number of followings

<table>
<thead>
<tr>
<th>Gender (persons in sample)</th>
<th>Total number of followings</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (945)</td>
<td>770736</td>
<td>434</td>
<td>16</td>
<td>85500</td>
</tr>
<tr>
<td>Women (1234)</td>
<td>755261</td>
<td>418</td>
<td>14</td>
<td>8676</td>
</tr>
</tbody>
</table>

\textsuperscript{10}This particular individual had recently deleted many of her tweets but the API Twitter collection program has collected 2534 tweets from collection point prior to June 2016.
Table 6.5 Summary data on the number of followers

<table>
<thead>
<tr>
<th>Gender (persons in sample)</th>
<th>Total number of followers</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (946)</td>
<td>1471879</td>
<td>398</td>
<td>10</td>
<td>220823</td>
</tr>
<tr>
<td>Women (1234)</td>
<td>865853</td>
<td>396</td>
<td>9</td>
<td>19800</td>
</tr>
</tbody>
</table>

Table 6.6 Summary data on the number of likes

<table>
<thead>
<tr>
<th>Gender (persons in sample)</th>
<th>Total number of likes</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (941)</td>
<td>2226707</td>
<td>884</td>
<td>1</td>
<td>102000</td>
</tr>
<tr>
<td>Women (1224)</td>
<td>3355127</td>
<td>1032</td>
<td>1</td>
<td>43300</td>
</tr>
</tbody>
</table>

In Table 6.3, the total number of tweets posted by the women is 1.4 times higher than that of the men which coincide with the higher median of tweeting activity. Table 6.4 portrays the number of ‘followings’ (i.e. the number of other Twitter users the people in the sample follow) and does not show a noticeable difference between men and women, except for the maximum number of ‘followings’ which is found in the male group with the maximum number almost ten times as high as the maximum found in the woman group. One man is an outlier who follows that many others and is followed by 152,000 in return. However, his profile gives no indication of why he follows that many people. The number of followers (other Twitter users following individuals in the sample) does not differ substantially in the total number of followers or of the median (Table 6.5) number across the genders. The maximum number of followers for men, however, is considerably higher and is possibly due to one man being the guitarist in a band which increased his number of followers tremendously. As demonstrated in Table 6.6, women are overall more probable of ‘liking’ someone else’s tweets. However, the maximum amount of ‘likes’ is done by a man and is more than twice as high as the highest ‘likes’ for women. Ten women and five men have not ‘liked’ any post on Twitter.

So in summary, this gender breakdown of activity suggests that women tweet and ‘like’ more but that there is very little gender difference in the median ‘followings’ and ‘followers’. This pattern of activity has been noted in other Twitter research. The Pew Research Centre, for example, who report on Mobile Messaging, and Social Media use in the USA found that women are more active on Twitter than men but that there are no differences in race, ethnicity or socio-economic position (Lenhart, Purcell, Smith, & Zickuhr, 2014). This same report found that women are more
likely to share content on Twitter compared to men (Lenhart et al., 2014). The observations made here and those from wider research suggest that gender is an important factor which may influence Twitter activity and smoking outputs, but in different ways. For this reason, additional results as summarised below, are stratified according to gender where that would be appropriate.

6.1.3 Twitter activity according to age

The presumed age of the sample members was derived from either their Twitter profiles or usernames containing a birthyear (which was transformed into their age on January 1st 2016). To examine how Twitter activity varies according to age, and for ease of analysis, the specific ages were categorised into four age groups: 18 and under, 19-21, 22-24, and 25 and over. This cut off point was selected out of convenience but, coincides approximately with different stages in education: 18 and under (high school and below), 19-21 (college or university), 22-24 (early employment or finalising education) and 25 and over (employment). Presumably, the tweets from the sample members with an age indication are made when they were in a different age category (i.e. tweets made before 2013). In the results, the age indication shows the tweets from that age group or possibly when the people were in a younger age group. Therefore, the results of this study are discussed as indications of the younger and older sample members.

According to Figure 6.2 below presenting the median for each type of Twitter activity per age group, the 19 to 21-years-old group tweet the most. They are also the age group where they follow more people and are themselves being followed the most. Interestingly, ‘liking’ someone else’s tweets has a gradient in which it becomes less prevalent with age. Another noteworthy finding is that for young people under the age of 21, it is more common to have more ‘followers’ than they ‘follow’, and this is reversed for the people aged 22 and above which means the older age groups follow more Twitter users that do not follow them back in return. The results might indicate that the way Twitter is used changes with age. When people get older, they mainly use Twitter as a form of information gathering rather than information spreading or socialising (D. J. Hughes et al., 2012; P. R. Johnson & Yang, 2009). This premise would also explain the less ‘likes’ in older age groups, and the shift towards the production of considerable fewer tweets compared to the younger
groups. The variance in Twitter activity is minimal, but further differences between age group should be considered as related to both smoking and the general Twitter activity.

Figure 6.2 Distribution of Twitter engagement per age group

6.1.4 Smoking status characteristics

An assumed smoking status was derived for each of the sample members through an analysis of the content of any tweets they posted that contained text relating to smoking. The smoking status was assessed by reading and evaluating, in chronological order, the smoking-related tweets for each individual sample member. The smoking status was based on the process that could be derived from their smoking-related tweets and were classified at the end of the data collection period. For example, if an individual only tweeted “coffee and a cigarette”, this resulted in that person being classified as a smoker. However, if another person first tweeted about having a cigarette but later posted that they “quit smoking two weeks ago” they were classified as a quitter as this person progresses from a smoker into a quitter. Likewise, when people tweeted about ‘having quit smoking’ but later on, tweeted about ‘having a cigarette’, they were classified as relapsers. The categorising of
smoking status across individuals is explained in more detail in the methods chapter (section 4.5.2 of Chapter 4). Table 6.7 presents how the sample is divided according to their smoking status. The highest number of young people in this sample are non-smokers (838; 38.4%) followed by current smokers (693; 31.8%). There were 403 (18.5%) quitters in the group, and the least prevalent group were the relapsers (246; 11.3%). It is not the intention of this analysis to compare these proportions with equivalent proportions across the general or youth population for Wales. Rather this study focuses on their smoking status as the sample members portray it on Twitter.

Table 6.7 Smoking status across the study sample. Row percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Number of people</th>
<th>Non-smoker (%)</th>
<th>Smoker (%)</th>
<th>Quitter (%)</th>
<th>Relapser (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2180)</td>
<td>838 (38.4)</td>
<td>693 (31.8)</td>
<td>403 (18.5)</td>
<td>246 (11.3)</td>
</tr>
</tbody>
</table>

6.1.5 Smoking status according to age and gender

Smoking status was analysed according to gender and age. The summary data of the frequency distribution of smoking status and gender are provided in Table 6.8. The table shows only small differences in the prevalence per smoking status across genders. This notion is confirmed by the chi-square test as smoking status is not statistically associated with gender in this study ($X^2= 4.17$, $df= 3$, $p=0.24$).

Table 6.8 Smoking status among men and women. Row percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Gender</th>
<th>Total (2180)</th>
<th>Non-smoker (%)</th>
<th>Smoker (%)</th>
<th>Quitter (%)</th>
<th>Relapser (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (946)</td>
<td>157 (37.3)</td>
<td>142 (33.6)</td>
<td>74 (17.5)</td>
<td>50 (11.8)</td>
<td></td>
</tr>
<tr>
<td>Women (1234)</td>
<td>264 (42.0)</td>
<td>176 (28.0)</td>
<td>110 (17.5)</td>
<td>78 (12.4)</td>
<td></td>
</tr>
</tbody>
</table>

The table of smoking status per age group is shown in Table 6.9 and exhibits some interesting results; the non-smokers is the largest in the sample (421) and is followed by the smoker group (318). However, the smokers do not have the highest count in the ‘25 and over’ category as was expected from the results of the Welsh Health Survey (2014) described in section 1.1.3 in Chapter 1. There is a gradient for quitters as the number of quitters goes up with each age group. Furthermore, apart from the
youngest age group, relapsing becomes less common with age which together with
the results from the quitters seems to indicate that when people get older, their
quitting attempts become more successful. A chi-square test ($\alpha=0.05$) reveals that
there is indeed an association between age group and smoking status ($X^2= 56.36$, $df= 9$, $p<0.01$).

**Table 6.9 Distribution of smoking status per age group. Row percentages are given between
brackets (%).**

<table>
<thead>
<tr>
<th>Age groups</th>
<th>Total (597)</th>
<th>Non-smoker (%)</th>
<th>Smoker (%)</th>
<th>Quitter (%)</th>
<th>Relapser (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 and under (77)</td>
<td>34 (44.2)</td>
<td>28 (36.4)</td>
<td>7 (9.1)</td>
<td>8 (10.4)</td>
<td></td>
</tr>
<tr>
<td>19-21 (222)</td>
<td>91 (41.0)</td>
<td>88 (39.6)</td>
<td>19 (8.6)</td>
<td>24 (10.8)</td>
<td></td>
</tr>
<tr>
<td>22-24 (180)</td>
<td>56 (31.1)</td>
<td>66 (36.7)</td>
<td>36 (20.0)</td>
<td>22 (12.2)</td>
<td></td>
</tr>
<tr>
<td>25 and over (118)</td>
<td>32 (27.1)</td>
<td>28 (23.7)</td>
<td>46 (39.0)</td>
<td>12 (10.2)</td>
<td></td>
</tr>
</tbody>
</table>

**6.2 Mapping the sample**

Attention now turns to the use of the geolocation coordinates extracted from the
Twitter archives of the sample members. This will provide an insight into the place
of residence of these young people and addresses the objectives concerning the
(geographical and demographic) reach of The Filter Wales.

The study sample consisted of a total of 2180 young people who generated over 4.8
million tweets during the period February 2007 to June 2016. Out of these tweets,
39,4775 (7.9%) were accompanied by geolocation coordinates. These locations
are stored as eight-digit latitude and longitude coordinates which means that a
person’s location can be resolved to an accuracy of 10 square metres (M. Graham
et al., 2014).

To illustrate the totality of this information, Figure 6.3 exemplifies all of the derived
geolocation points placed on a world map. It should be noted that there may be
several tweets posted from the same location which is represented by a single dot.
The map, therefore, does not represent the density of tweets across geography but
instead portrays the geographical extent of tweeting activity. The map represents the
group of people who have come into contact with The Filter Wales because every
individual has tweeted, at some point in time, about smoking from within the area of
Wales (or very close to its borders).
Figure 6.3 Distribution of Twitter geolocation coordinates on a world map.
In addition to the many tweets generated in Europe, much Twitter activity originates from English-speaking nations such as the USA and New Zealand. The possible explanations of why they are tweeting from all around the world are plentiful. They could have been on holiday to these locations, taking a gap year, undertaking international internships, or engaging in overseas volunteering. However, as the sample members are not contacted, there is no concrete evidence to support these potential reasons for tweeting outside of the UK.

6.2.1 Tweets across the UK

Unsurprisingly, most tweets are sent from the UK and the map in Figure 6.4 illustrates the distribution of geolocated tweets within the UK and Ireland. This map reveals that most of those tweets are sent from the major UK cities (e.g. London, Birmingham, and Liverpool) and may represent locations where members of the sample are students, visiting relatives or they may have jobs in these locations. Interestingly, Figure 6.4 picks up ‘trails’ of tweets which coincide directly with major motorways and train lines. These trails imply that the people in the sample are actively tweeting while taking a train or car journey.

Taking a closer look at the geolocation coordinates plotted across Wales (Figure 6.5), it becomes apparent that, by far, most tweets are posted from the urban areas of South Wales such as Cardiff, Newport and Swansea. Moreover, a clustering of dots appears on the coastline of North Wales close to places such as Conwy and Rhyl. The map shows the Lower Super Output Area (LSOA) boundaries found within Wales (see section 4.5.5 of Chapter 4 for a definition of LSOAs), and while there is less activity in some of the more sparsely populated rural areas, almost all LSOAs have at least one tweet being sent from within their boundaries. Distinct patterns emerge of strings of geolocated tweets moving along transport lines in the cities or the more rural areas (Figure 6.5) which are similar to the patterns from Figure 6.4 and indicate that young people are tweeting during their everyday commute.
Figure 6.4 Distribution of Twitter geolocation coordinates within the United Kingdom and Ireland
Figure 6.5 Distribution of Twitter geolocation coordinates within Wales with an underlay of LSOA boundaries.
6.2.2 Using the geolocation information to reveal likely place of residence

The grid coordinates from the previous section are linked to the LSOAs to explore the geography of the sample tweets and Twitter users. Any user may tweet from several different locations, and so the variation in these areas needs to be studied in more depth to find the LSOA where the individual probably resides. Here the assumption is that, for each young person, the most prevalent LSOA was probably their place of residence (see Chapter 4 section 4.5.5 for more details). Johnson & Yang (2014) questioned Twitter users about their motives for using Twitter and uncovered that people often used it in times of inactiveness such as during a commute or at home in front of the TV.

A limit was set to at least five coordinates within any one English and Welsh LSOAs to ensure there was enough evidence to determine the likely place of residence. Out of the total 1678 people with geolocation attached to their tweets, 428 had 4 or fewer LSOAs to their name and were eliminated from further analysis. Once the predominant LSOA was assigned to each individual, the ones that ended up with the allocation of an English LSOA (199) were then eliminated from the table as they are not assumed inhabitants of Wales which was necessary for the next step in the analysis. A sample of 1051 people with a Welsh LSOA was left for further analysis.

Using this assumed residence information, Figure 6.6 shows the distribution of the sample members. The darker areas indicate places with more members and the lighter areas have fewer sample members. While several areas (especially rural areas) do not have any sample members, the map shows that the study sample is dispersed (regarding residential location) across the country. Most of the individuals are from South Wales where population densities are high, and these areas also contain universities, colleges and work opportunities which makes their presence more likely. Slightly higher numbers of individuals are also visible around western mid-Wales and especially close to Aberystwyth which is a university city. Further in the north, there are some residential areas where clusters of Twitter users can be found.

One suggestion is that the awareness of The Filter Wales has increased the number of smoking-related tweets in the areas around Cardiff (ASH Wales’s office) because the outreach activities surrounding the campaign are much more visible. However, the social media team works separately from The Filter Wales youth outreach events. Therefore, it is likely that the awareness of The Filter Wales’ existence increases the presence of smoking tweets for young people, but it cannot be measured in this study. However, the outreach
events do not overlap with the assumed place of residence of the sample members as the team has been to Wrexham in the North-East of Wales and Carmarthen west of Swansea, but these areas show relatively few sample members. Moreover, The Filter Wales has not been in Aberystwyth, but 31 individuals in this study have a place of residence in that Aberystwyth area.

To have a closer look at those areas where there is a relatively high density of engagement, Figures 6.7a to 6.7d depict close-ups of such areas in the south, mid-west and north of Wales. First, the Swansea area which is Figure 6.7a reveals that sample members are living throughout this region in various densities but are mainly found close to the urban centres
of Swansea, Llanelli and Bridgend. Similarly, in Figure 6.7b, the LSOAs with a high number of sample members are located in Cardiff and to a lesser degree near the other urban areas of South Wales. It is evident from Figure 6.7c that Aberystwyth and the surrounding rural areas are less populated as the LSOAs are larger than in South Wales. Most of the users here are found in the student town of Aberystwyth with its relatively large student population. Lastly, Figure 6.7d illustrates the distribution of the sample in the Northern part of Wales. The towns and cities in the north (Rhyl, Conwy and Bangor) are not as densely populated with sample members. There is a higher density in the eastern edge of Rhyl, but absolute numbers are not comparable with the urban areas in the south.

The overall conclusion is that the allocation of youth in the sample is distributed across most of Wales, but the majority lives in the more densely populated urban areas in the south. This information on the assumed place of residence of the individuals demonstrates an important finding that the Filter Twitter campaign has reached young people from all across Wales.

Figure 6.7 Overview map of specific areas for close-up Figures 6.7a-d
Figure 6.7a Count of sample members per LSOAs in the Swansea area (South Wales)

Figure 6.7b Count of sample members per LSOAs in the Cardiff area (South Wales)
6.2.3 Smoking Twitter activity per local authority

The health services in Wales are arranged in 22 local authorities (LAs). Many LAs have local tobacco action groups that implement the Tobacco Control Action plan throughout Wales with a target to reduce smoking prevalence to 16% by 2020 (Ministry for Health
and Social Services, The Welsh Government, & Ministry for Health and Social Services, 2012). An important factor in decreasing smoking prevalence is the tobacco control services focusing on youth. The socio-economic locational variances and the different population build-up between LAs are why responsibility is given to LAs to develop their own local comprehensive plan to deliver tobacco control services (Ministry for Health and Social Services et al., 2012). The Tobacco Control Action Plan does not specifically target youth populations, but from a local authority perspective, it might be useful to know the extent to which younger groups are engaging in smoking-related Twitter interactions. To examine this in more depth, the level of engagement was determined by comparing smoking-related Twitter activity against the percentage of young people in each LA. The age range of this youth population was defined as aged between 11-25. This is because 99% of smokers started between these ages (ASH, 2015c) and The Filter Wales utilises the same target group upon which to focus their activities.

Wales consists of 22 Local Authorities with a youth population varying from 15.11% (Isle of Anglesey) to 37.2% (Ceredigion). Not all smoking-related tweets could be used in this analysis as there needed to be an indication of the place of residence of the sample member. The 8338 (49.9%) tweets from the sample population with an assumed place of residence were assessed for the correlation between smoking-related tweets and the proportion of youth population within each LA.

The scatterplot (Figure 6.8) summarises the results with most LAs following a positive trend, in which LAs with a higher youth population have more smoking-related tweets in the dataset. A Pearson’s \( r (\alpha=0.05) \) was computed to test the relation between the percentage of the youth population in a Local Authority and the number of smoking-related tweets from this sample. There was a correlation between the two variables \( (r=0.516, n=22, p=0.014) \) which indicates a significant association. However, there are a few outliers visible in the scatterplot such as Ceredigion, Carmarthenshire and Gwynedd. The LA of Ceredigion consists of a high population of 11 to 25-year-olds (37.2%), but the number of smoking-related tweets found in this sample is low (269 tweets). The high youth population in the Ceredigion originates from the Aberystwyth University situated in that LA. One of the LSOAs in Aberystwyth has a youth population of 83.1% which skewed the results for the entire LA. The LA of Carmarthenshire is located to the west of Swansea in the Southwest of Wales and includes the town of Llanelli. This LA has a relatively high number of smoking tweets (545) compared to the percentage of the youth population (16.42%). The last noteworthy outlier is Gwynedd. This LA has quite a high prevalence of
youth (20.25%), but there are a low number of smoking-related tweets posted from this area (98 tweets) associated with this study. A possible explanation is that the Gwynedd area has the highest number of Welsh speakers (Office for National Statistics, 2014a) but tweets in Welsh were not included in the smoking-related dataset.

![Figure 6.8](image)

**Figure 6.8** Scatterplot of the number of smoking-related tweets per percentage of youth population per local authority in Wales.

The smoking-related tweets are divided over the LAs, and there is a positive correlation between the youth population per LA and the number of smoking-related tweets indicating that if there are more young people in that area, there are a higher number of tweets. This is, of course, dependant on the reach of The Filter Wales Twitter feed as only the smoking tweets from people in contact with The Filter Wales were considered in this study.

Figures 6.7a to 6.7d and 6.8 demonstrate how The Filter Wales has been in contact with young people from all over Wales which includes at least several people from each LA and that the number of smoking tweets correlated with the proportion of youth in the LA.
6.3 Combining geolocation with place characteristics

Another way to look at the influence of place is through deprivation and levels of rurality or urbanicity. In the previous section, the geolocation coordinates (where available) were used to determine the likely LSOA of residence of sample members. Connecting the people in the sample to an LSOA has the advantage of connecting the place of residence to other geographical components. This section, therefore, focuses on understanding the deprivation context and the rural/urban characteristics of the place of residence by analysing them across gender and smoking status. The understanding of neighbourhood associations with smoking behaviour can help target smoking cessation programs, and for this thesis, it contributes to one of the main objectives of the research which is to understand more about the reach of The Filter Wales regarding engagement across more deprived areas and more remote locations.

6.3.1 Neighbourhood deprivation

Neighbourhood deprivation is summarised using the 2014 Welsh Index of Multiple Deprivation (Welsh Government, 2014) and is based on eight domains of socio-economic advantage/disadvantage covering income, employment, health, education, access to services, community safety, physical environment, and housing. The Indices of Multiple Deprivation in Wales (WIMD) are calculated in a way that the most affluent neighbourhoods receive a high total score.

Figure 6.9, adapted from the Welsh Health Survey 2014, reveals the smoking gradient across the deprivation quintiles, where smoking prevalence is lowest in the most affluent area and increases as deprivation increases. The prevalence of adult smoking in the most deprived areas (29%) was considerably higher than in the most affluent areas (11%) (Statistics For Wales, 2015).
Interestingly, historical data on the deprivation-smoking link since 2003/4 indicate that the disparities between deprivation levels have always been visible, but the gap between the more and less deprived areas has reduced (Figure 6.10). The most deprived areas have the most extensive decline in smoking prevalence but in 2015, still report a smoking prevalence which is twice as high as in the least deprived area, as evidenced in the line chart below (Statistics For Wales, 2015).

Figure 6.9 Adults who report smoking by deprivation quintile (source: Welsh Health Survey, 2014)

Figure 6.10 Prevalence adult smoking by Welsh Index of Multiple deprivation quintiles (source: Welsh Health Survey 2014)
Using the WIMD in this study provides valuable information on the place context of the young people in the sample. The graphs, outlined above, illustrate the clear link between deprivation and higher smoking rates. It is, therefore, important that The Filter Twitter campaign reaches across all deprivation profiles and arguably, has a greater presence in more deprived areas.

Tables 6.10 and 6.11 show the distribution of the sample of young people according to the quintiles of deprivation across the LSOAs. The level of deprivation is derived from the national quintiles, so it is to be expected that approximately 20% of the sample population would be found in each of them if there was no difference between the deprivation profile of the sample and the population of Wales.

Table 6.10. Neighbourhood descriptive statistics across quintiles of WIMD ranks in Wales. Column percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Quintiles of deprivation based on rank</th>
<th>Sample 1051 (%)</th>
<th>Men 423 (%)</th>
<th>Women 629 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (1 – 382) Most deprived</td>
<td>241 (22.9)</td>
<td>105 (24.8)</td>
<td>136 (21.6)</td>
</tr>
<tr>
<td>2 (383– 764)</td>
<td>189 (18.0)</td>
<td>77 (18.2)</td>
<td>112 (17.8)</td>
</tr>
<tr>
<td>3 (765 – 1146)</td>
<td>207 (19.7)</td>
<td>87 (20.6)</td>
<td>120 (19.1)</td>
</tr>
<tr>
<td>4 (1147- 1528)</td>
<td>196 (18.7)</td>
<td>68 (16.3)</td>
<td>128 (20.3)</td>
</tr>
<tr>
<td>5 (1529 - 1909) Least deprived</td>
<td>218 (20.7)</td>
<td>85 (20.1)</td>
<td>133 (21.1)</td>
</tr>
</tbody>
</table>

Table 6.10 provides the detail on the deprivation profile of the sample and offers the distribution according to gender. From here, it can be understood that the sample population reflects a similar deprivation profile to the general Welsh population, with only a small increase in this sample from the most deprived area. The table demonstrates that the distribution of men in the sample across different deprivation quintiles shows that nearly a quarter lives in the most deprived area. If that is compared with the women, that difference in representation per quintile is not as clear. There appears to be equality in engagement with the Filter Twitter feed according to deprivation as the association measured by the Chi-square test ($\alpha=0.05$) is not significant between gender and quintiles of deprivation level ($X^2= 3.61$, $df= 4$, $p=0.46$). It would, therefore, appear that the Filter has a similar reach across areas according to deprivation profile and for both genders.

Table 6.11 presents the smoking status of the young people in the sample as distributed across the ranked quintiles of Welsh multiple deprivation. The Chi-square test ($\alpha=0.05$) reports a significant association between smoking status and quintiles of neighbourhood
deprivation ($X^2= 24.01$, $df= 12$, $p=0.02$). Out of the non-smokers, most are from the least deprived communities as are the relapsers. The smokers tend to be equally distributed among the quintiles. Interestingly, the quitters are more dominant in the more deprived areas. This result is interesting and challenges common rhetoric which suggests that people in more deprived areas have more difficulty quitting smoking (e.g. Hiscock et al., 2012). Most likely in this sample, the smoking status is produced from the online expression of smoking which favour success stories about having quit smoking over posts of failure in the form of relapsing back into smoking. Therefore, the results likely show a variance of the Twitter style rather than actual smoking status between deprivation levels.

Table 6.11 Smoking status of the Twitter users per WIMD scores in Wales. Column percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Quintiles of deprivation</th>
<th>Non-smokers 421 (%)</th>
<th>Smokers 318 (%)</th>
<th>Quitters 184 (%)</th>
<th>Relapsers 128 (%)</th>
<th>Total 1051 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (most deprived)</td>
<td>78 (18.5)</td>
<td>68 (21.4)</td>
<td>44 (23.9)</td>
<td>28 (21.9)</td>
<td>241 (22.9)</td>
</tr>
<tr>
<td>2</td>
<td>61 (14.5)</td>
<td>65 (20.4)</td>
<td>45 (24.5)</td>
<td>25 (19.5)</td>
<td>189 (18.0)</td>
</tr>
<tr>
<td>3</td>
<td>83 (19.7)</td>
<td>66 (20.8)</td>
<td>39 (21.2)</td>
<td>19 (14.8)</td>
<td>207 (19.7)</td>
</tr>
<tr>
<td>4</td>
<td>84 (20.0)</td>
<td>51 (16.0)</td>
<td>30 (16.3)</td>
<td>24 (18.8)</td>
<td>196 (18.7)</td>
</tr>
<tr>
<td>5 (least deprived)</td>
<td>115 (27.3)</td>
<td>68 (21.4)</td>
<td>26 (14.1)</td>
<td>32 (25.0)</td>
<td>218 (20.7)</td>
</tr>
</tbody>
</table>

6.3.2 Urban/rural classification

The 2011 rural/urban classification (DEFRA, 2011) is produced by the UK government and classifies LSOAs into ten categories based on the population of the area the LSOA is a part of and the surrounding neighbourhood. An area is classified as urban if it has a population of 10,000 or more. The urban options are; Major Conurbation, Minor Conurbation, City and Town, City and Town in a Sparse Setting. The LSOAs in areas with less than 10,000 inhabitants were considered rural and rural options are; Town and Fringe, Town and Fringe in a Sparse Setting Village, Village in a Sparse Setting Hamlets, and Isolated Dwellings Hamlets, and Isolated Dwellings in a Sparse Setting. There is an extent of the difference between smoking behaviour in the cities compared to the countryside (Gartner, Farewell, Roach, & Dunstan, 2011; Law & Morris, 1998). Gartner et al. (2011) and Law & Morris (1998) found that people in their samples smoked most in rural areas and the more deprived neighbourhoods in cities which leads to higher smoking-related mortality in those areas. Moreover, rurality induces a lack of smoking cessation resources and interventions (Hutcheson et al., 2008).
The assumed LSOAs of residence of the sample members are classified into their type of rural/urban characteristics as defined by DEFRA’s classification system and are shown in the table below (Table 6.12). The first important observation is that the sample has a quite substantial overrepresentation of people from LSOAs from an ‘Urban city or town’ (79.8%) compared to the proportion of LSOAs that are classified as urban city or town in Wales (66.4%). The high proportion of the sample living in an urban city or town was expected when looking at Figure 6.7 as they live close to universities, colleges and work. There is no apparent difference between the different types of areas which are identified by the Chi-square test (α=0.05) not being significant ($X^2 = 3.51, df = 5, p = 0.62$).

Table 6.12. Neighbourhood descriptive statistics with Rural/Urban classification in Wales. Column percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Rural/urban classification</th>
<th>Welsh LSOAs 1909 (%)</th>
<th>Sample 1051 (%)</th>
<th>Men 423 (%)</th>
<th>Women 629 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural town and fringe</td>
<td>252 (13.2)</td>
<td>93 (8.8)</td>
<td>31 (7.3)</td>
<td>62 (9.9)</td>
</tr>
<tr>
<td>Rural town and fringe in a sparse setting</td>
<td>78 (4.1)</td>
<td>30 (2.9)</td>
<td>10 (2.4)</td>
<td>20 (3.2)</td>
</tr>
<tr>
<td>Rural village and dispersed</td>
<td>129 (6.8)</td>
<td>31 (2.9)</td>
<td>14 (3.3)</td>
<td>17 (2.7)</td>
</tr>
<tr>
<td>Rural village and dispersed in a sparse setting</td>
<td>147 (7.7)</td>
<td>31 (2.9)</td>
<td>11 (2.6)</td>
<td>20 (3.2)</td>
</tr>
<tr>
<td>Urban city and town</td>
<td>1268 (66.4)</td>
<td>839 (79.8)</td>
<td>345 (81.6)</td>
<td>494 (78.7)</td>
</tr>
<tr>
<td>Urban city and town in a sparse setting</td>
<td>35 (1.8)</td>
<td>27 (2.6)</td>
<td>12 (2.8)</td>
<td>15 (2.4)</td>
</tr>
</tbody>
</table>

Table 6.13 reports the dissemination of smoking status according to a simple two-fold urban/rural classification system. This is created by defining the two urban option as ‘urban’ and the four rural options as ‘rural’. For this table (6.13), row percentages are given as the distribution of sample members was so unevenly distributed that the table would not show any distinctions except that the vast majority within each smoking status lives in an urban location. Instead, the table shows how smokers are more from urban areas compared to the other types of smoking status. Relapsers have a comparably high percentage of the group living in more rural areas. Similar to the gender division, there is no significant (α=0.05) association between the urban/rural divide and smoking status ($X^2 = 6.03, df = 3, p = 0.11$).
Table 6.13 Smoking status of the Twitter users per Urban/Rural Divide* in Wales. Row percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Urban/rural classification</th>
<th>Non-smoker</th>
<th>Smoker</th>
<th>Quitter</th>
<th>Relapser</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban 866 (100%)</td>
<td>421 (%)</td>
<td>318 (%)</td>
<td>184 (%)</td>
<td>128 (%)</td>
</tr>
<tr>
<td>Rural 185 (100%)</td>
<td>338 (39.0%)</td>
<td>275 (31.8%)</td>
<td>152 (17.6%)</td>
<td>101 (11.7%)</td>
</tr>
</tbody>
</table>

*the rural/urban classification was dichotomised: Urban city and town (in a sparse setting) were classified as 'urban' and all others as 'rural'.

To sum up, the sample population was almost equally divided among the categories of multiple neighbourhood deprivation, and by far most young people lived in an urban city or town, but this was not associated with gender. Smoking status, on the other hand, did have an association with neighbourhood deprivation but not with urban/rural living. Even though The Filter Wales have reached people from all across Wales; they did not have an extended reach in more deprived or rural areas which would be the types of places in highest need of the services of a youth dedicated smoking cessation organisation.

6.4 Concluding remarks

One of the objectives of the chapter was to determine the gender, age, smoking status, and place of residence of the young people that have been in contact with The Filter Twitter element by gathering profile data and geolocation coordinates from the Twitter users. Gender is a common variable in Twitter research, and many API Twitter programs can derive gender automatically. As a result, many studies have used gender in their analysis (e.g. Hughes, Rowe, Batey, & Lee, 2012; Mislove, Lehmann, Ahn, Onnela, & Rosenquist, 2011; Myslín, Zhu, Chapman, & Conway, 2013). Their general conclusion is that women are more active on Twitter, but the variance is minimal as is similar in this study. Age was harder to derive from the Twitter profiles, and only 27.4% of the sample members gave an indication of age. The results of this chapter indicate some importance of age combined with Twitter activity as older sample members used Twitter more for information gathering than socialising. These two uses are the main motives for using Twitter according to Johnson & Yang (2009) and Hughes et al. (2012).

Information on youth smoking status is generally derived from questionnaires about smoking habits which can be associated with other variables such as gender, age and place of residence (such as in the HSBC and Welsh Health Survey). In this study, smoking status was not derived from direct questioning. Instead, through necessity, it was derived from an analysis of the text contained in the sample members’ tweets. The original aims of such
tweets are not to represent actual smoking status but are used as a means for young people to communicate interesting things to their followers (Hutto, Yardi, & Gilbert, 2013). Although there is little opportunity to verify the assumed smoking status constructed in this way, its derivation which is less intrusive than direct face-to-face questioning may well improve the validity of the information. The analysis presented above estimated that there were mostly non-smokers in the sample (38.4%), followed by smokers (31.8%), quitters (18.5%) and lastly relapers (11.3%).

For the second objective was to map the reach of the Filter and therefore determine the likely place of residence which was possible for 1051 (48.2%) sample members. The members derived from across the whole of Wales but most resided in the south close to Cardiff and Swansea. When analysed at Local Authority level, the number of smoking-related tweets from the sample members with a place of residence indication was positively correlated with the percentage of the young people in the area (11-25 years of age). Interestingly sample members were detected in all the Local Authorities across Wales, indicating that the reach of The Filter Wales Twitter feed goes beyond the Cardiff hub of activities and suggests that it may engage with hard-to-reach young people in Wales. The deprivation analysis revealed that most non-smokers and relapers are from the most affluent areas, the quitters are from the most deprived, and there was no clear difference between smoking status and urban/rural classification in this sample.

Overall, this study strengthens the perception that Twitter feeds contain a vast number of information about their Twitter users. Moreover, the geolocation attached to the tweets helped contextualise these results across location and types of place. In the next three chapters, the focus shifts from the demographic and geographical characteristics of the individuals posting the tweets and places greater attention on smoking behaviour as revealed via different types of content analysis of the text in each tweet. Chapter 7 provides a quantitative content analysis in which comparisons are made between the different smoking products and the different type of smoking actions. Additionally, this chapter demonstrates the variations in the timing of tweets and the sentiment attached to them to broaden understanding of the social meaning of smoking, a topic which is further explored in Chapters 8 and 9.
Chapter 7. Uncovering the content of the smoking tweets

The previous chapter provided a summary of the information that could be derived from the research sample. This chapter turns to the content of the tweets which have been written by the individuals in the sample. The focus centres on the production of a quantitative summary and provides results for two objectives:

- to reveal the engagement with the Twitter feed, including differences in the subject content of tweets posted
- to establish the temporal and sentimental qualities of the smoking-related tweets

These objectives facilitate a deeper understanding of what place smoking holds in the everyday lives of the sample members. This is done by looking at the number of smoking-related tweets posted by the sample population, by plotting the frequency of tweets on timeframes to establish when people are posting smoking-related content, and by analysing, in more depth, the positive, neutral or negative sentiment characteristics of the text content. Undertaking this analysis is essential because the content of the smoking referenced tweets are a key indicator of how the sample members perceive their smoking behaviour and that of others.

The first section (7.1) of this chapter focuses on the overall content of the tweets characterised by the type of smoking product, the person performing the smoking, and the type of smoking activity (see Chapter 4, Section 4.5.3 for a summary of this data preparation). The chapter moves on to examine the Twitter activity per smoking product over time with hours, weekdays, and months in comparison to general Twitter activity (section 7.2). This section further describes the tobacco-related Twitter activity on the yearly timeline and presents an exploratory graph of the tobacco smoking action to clarify the peaks at specific times. Following the temporal qualities, the sentimental qualities are assessed in section 7.3. The tweets are analysed according to their positive and negative sentiment scores. This section states how the sentiment scores are calculated and the meaning of these scores. For the final part of the chapter (section 7.4), the sentiment scores of the smoking-related tweets are placed on timelines to uncover when the sample members are most positive and negative about certain smoking products.

7.1 Revealing the content of tweets

The methods to create the table of smoking referenced data are described in Chapter 4 section 4.5, but to outline briefly: 4.8 million tweets from the sample members were uploaded unto the ORACLE database server, and through this program all the rows (i.e.
tweets) with a smoking reference into a separate database were exported. The resultant 16,688 number of tweets were read individually by the researcher and systematically categorized according to the type of smoking product referenced in the tweet (i.e. tobacco, marijuana, e-cigarettes, and shisha), the person performing the action i.e. the originator ‘First Person’ (FP), anyone else ‘not First Person’ (Not FP), and the smoking action which is described in the text of the tweet (i.e. smoking, desiring to smoke, thinking of quitting, and quitting smoking). For example, the tweet “I reeaalльy want a cigarette” was categorised as ‘tobacco’, ‘FP’, and ‘desiring to smoke’. Chapter 4 section 4.5.3 provides a more detailed description of how the smoking-related tweets were collected and how the tweet content was associated with specific characteristics. The number of tweets that could be used in the content analysis is 10,391 (out of the total of 16,688 tweets).

7.1.1 Who is smoking what?

Similar research in this area has stated that most tweets contain a reference to the individual posting the tweet and that the activity relates to smoking tobacco or marijuana cigarettes (Cavazos-Rehg et al., 2015; Murnane & Counts, 2014; Myslín et al., 2013; Prier et al., 2011; Leah Thompson, Rivara, & Whitehill, 2015). This study starts by examining the content for this study population to establish if these individuals tweet similar content. Table 7.1 below lists the tweet categorisations and the number of tweets fitting those descriptions across all the smoking-related posts which were made between 2 April 2009 and 30 June 2016. The columns in this table refer to the smoking-related action described in the tweet. The rows list the type of smoking product, and each is further subdivided according to the ‘person’. Table 7.1 demonstrates clearly that almost three-quarter of the tweets has tobacco as their subject and that those tweets primarily focus on the individual who posted the tweet (FP). The most mentioned tobacco smoking activity is either ‘smoking’ followed by ‘quitting smoking’. This result can be explained by the content itself as tweets about smoking and having quit smoking can be posted many times; each time someone has a cigarette and the number of days they have quit. Another motive for posting this type of content is the validation these posts receive. The popularity of ‘quitting’–related content may indicate that the individual consecutively wants to explain their achievement in the quitting attempt, e.g. “it’s been 3 days since I quit” followed by “it’s been a week since I last had a fag”. The same could be argued for the ‘smoking’ activity.

However, when looking at many of those tweets in this tobacco group, the topic of the tweet was not smoking but instead something that happened while the FP smoked. An
example is; “Cool just stepped on a slug barefoot in my garden whilst having a fag”. These tweets were categorised as ‘Tobacco’, ‘FP’, and ‘smoking’ as it is the smoking-related message from the tweet that is important for the research presented here. This phenomenon of being ‘a side characteristic’ was specific to tobacco and was hardly seen in the tweets containing the other smoking products.

Table 7.1 Descriptive summary of tweets according to smoking product, action and who the action refers to.

<table>
<thead>
<tr>
<th>Smoking product (%)</th>
<th>Person</th>
<th>Number of tweets (%)</th>
<th>Smoking (%)</th>
<th>Desiring to smoke (%)</th>
<th>Thinking of quitting (%)</th>
<th>Quitting (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobacco 8050 (73.8)</td>
<td>FP</td>
<td>6850 (85.1%) of tobacco 1200 (14.9%) of tobacco</td>
<td>3055 (44.6)</td>
<td>902 (13.2)</td>
<td>1103 (16.1)</td>
<td>1790 (26.1)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td></td>
<td>985 (82.1)</td>
<td>68 (5.7)</td>
<td>89 (7.4)</td>
<td>58 (4.8)</td>
</tr>
<tr>
<td>Marijuana 1913 (20.5)</td>
<td>FP</td>
<td>1291 (67.5%) of marijuana 622 (32.5%) of marijuana</td>
<td>1016 (78.7)</td>
<td>184 (14.3)</td>
<td>35 (2.7)</td>
<td>56 (4.3)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td></td>
<td>581 (93.4)</td>
<td>23 (3.7)</td>
<td>15 (2.4)</td>
<td>3 (0.5)</td>
</tr>
<tr>
<td>E-cigarettes 266 (3.5)</td>
<td>FP</td>
<td>234 (88%) of e-cigarettes 32 (12%) of e-cigarettes</td>
<td>91 (38.9)</td>
<td>18 (7.7)</td>
<td>57 (24.4)</td>
<td>68 (29.1)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td></td>
<td>23 (71.9)</td>
<td>2 (6.3)</td>
<td>6 (18.8)</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>Shisha 162 (2.2)</td>
<td>FP</td>
<td>62 (38.3%) of shisha 100 (61.7%) of shisha</td>
<td>31 (50.0)</td>
<td>27 (43.5)</td>
<td>2 (3.2)</td>
<td>2 (3.2)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td></td>
<td>73 (73)</td>
<td>25 (25)</td>
<td>2 (2)</td>
<td></td>
</tr>
<tr>
<td>Total 10,391 (100)</td>
<td></td>
<td></td>
<td>5855 (56.3)</td>
<td>1249 (12)</td>
<td>1309 (12.6)</td>
<td>1978 (19)</td>
</tr>
</tbody>
</table>

Marijuana smoking was the second most popular type of smoking product in the content of the tweets from the sample (Table 7.1). Even though it is an illegal smoking product, the people in the sample seem to feel no obstruction tweeting about it, and their preferred topic is ‘smoking’ marijuana. Similar observations were made by Thompson, Rivara & Whitehill (2015) in their content analysis of marijuana-related tweets whereby 54.7% of tweets referred to the use (i.e. smoking) of marijuana. Posts with marijuana smoking content are perceived as favourable by their followers and are, therefore, posted more (Cavazos-Rehg et al., 2014). This system reinforces people to tweet about marijuana smoking which might have increased the number of marijuana-related tweets in this sample. Compared to tobacco use, ‘thinking of quitting marijuana’ and ‘quitting marijuana’ are not often tweeted. Amos et al. (2004) interviewed 15-16-year-old smokers about tobacco and marijuana, and none of them felt the need to quit smoking marijuana even when they were considering quitting tobacco smoking. This suggests that the two
activities are seen as separate, even though many marijuana smokers use tobacco in their marijuana cigarette.

In the e-cigarette category, unravelling ‘smoking activity’ posed some difficulty as e-cigarettes are advertised as a product for people to quit smoking tobacco. However, recent studies have illustrated that young people also take up smoking by starting with an e-cigarette (Eastwood et al., 2015; Goldstone et al., 2016; Ramo, Young-Wolff, & Prochaska, 2015) signifying that smoking an e-cigarette does not necessarily relate to quitting tobacco smoking. The systematic content analysis meant that the activity with the e-cigarette was paramount, resulting in tweets such as “Looks like today is the day I give up smoking and switch to e-cigs” were placed in the category of e-cigarette smoking and not ‘quitting tobacco’ because e-cigarettes are the topic of this tweet and not tobacco. There is no indication from the tweets in this section that the sample members started with e-cigarettes instead of tobacco. Instead, almost all ‘smoking e-cigarettes’ tweets contained information indicating a tobacco history such as; “I swear this electric fag makes me moody, its not the same”. As this is an important finding, policymakers and smoking cessation organisations should focus on e-cigarettes as a means to quit tobacco smoking and frame it as e-cigarettes being a less damaging alternative (Noland et al., 2016). However, e-cigarettes still pose health risks and, therefore, should not be the final stage but a step to becoming completely smoke-free.

Shisha is tweeted about the least out of the smoking products (as seen in Table 7.1). The bulk of shisha-related tweets are in the ‘not FP’ group as waterpipe smoking is mostly used in a social context (e.g. “@##### shisha tomorrow bra? Got all day”) and the shisha-related tweets would better suit the category ‘not just the First Person’. Shisha smoking is done as a social activity whereby people come together in a bar or just around a waterpipe. As mentioned earlier in Chapter 2 section 2.6, young people believe this is more entertaining than ‘normal’ tobacco smoking (Jawad et al., 2014). Even though it has the same components as tobacco or marijuana, is not seen as an addictive or harmful activity (Cobb et al., 2011; Martinasek et al., 2011). Only six tweets relate to quitting smoking shisha, and the few tweets in this groups are based on light-hearted reflective thoughts the young people have, e.g. “No more shisha till September @##### @##### how am I gonna cope ;);;”. The tweet content analysis here did not provide any clear indication that shisha smoking is prevalent among young people as only a few references mention shisha.

A large number of tweets (6297) were excluded from the quantitative content analysis as their content did not reference a smoking product, a person, and an action. These tweets
were placed in the ‘other’ category and essentially comprised of; too little information “yes weed”; smoking as the topic “how people find smoking attractive is beyond me”; promotions of the Filter “just won a beach break golden ticket @thefilter @cutfilms #overthemoon”; facts “smoking one cigarette takes about 11 minutes off of your lifespan”; new policies “car smoking ban, new law to protect children will come into action”; and things happening in smoking areas “what was that thing u whipped out in the smoking area?”.

7.1.2 Smoking and gender

This section discusses the smoking product content of the tweets according to gender (Tables 7.2 to 7.5). The first table on tobacco-related tweets (Table 7.2) does not show much variation between men and women and the way they tweet about tobacco use except that women post more frequently about ‘desire to smoke’ compared to men. Moreover, women are more likely to tweet about themselves ‘quitting smoking tobacco’ whereas men post more about themselves ‘smoking tobacco’.

<table>
<thead>
<tr>
<th>Tobacco (%)</th>
<th>Person</th>
<th>Number of tweets (%)</th>
<th>Smoking tobacco (%)</th>
<th>Desire to smoke tobacco (%)</th>
<th>Thinking of quitting tobacco (%)</th>
<th>Quitting tobacco (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men 3497 (43.4%)</td>
<td>FP</td>
<td>2971 (85%) of men</td>
<td>1401 (47.2)</td>
<td>329 (11.1)</td>
<td>511 (17.2)</td>
<td>730 (24.6)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>526 (15%) of men</td>
<td>442 (84.0)</td>
<td>20 (3.8)</td>
<td>35 (6.7)</td>
<td>29 (5.5)</td>
</tr>
<tr>
<td>Women 4553 (56.6%)</td>
<td>FP</td>
<td>3879(85.2%) of women</td>
<td>1654 (42.6)</td>
<td>573 (14.8)</td>
<td>592 (15.3)</td>
<td>1060 (27.3)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>674 (14.8%) of women</td>
<td>543 (80.6)</td>
<td>48 (7.1)</td>
<td>54 (8.0)</td>
<td>29 (4.3)</td>
</tr>
<tr>
<td>Total 8050 (100%)</td>
<td></td>
<td>4040 (50.2)</td>
<td>970 (12)</td>
<td>1192 (14.8)</td>
<td>1848 (23)</td>
<td></td>
</tr>
</tbody>
</table>

Marijuana has a more significant gender divide than tobacco as seen in Table 7.3, even though women tweeted more in the sample, the men tweeted the most about smoking marijuana. Studies on marijuana-related tweets report that men post more marijuana Twitter content but only by a small proportion (Cavazos-Rehg et al., 2015; Leah Thompson et al., 2015). However, the results presented here, clearly challenge this assumption. Furthermore, the tweets referring to the first person (i.e. the originator of the tweet) showed men tweeting about themselves was far more prevalent than for women, but the proportion of tweets for each smoking action in the tweet is almost equal.
Table 7.3 Descriptive summary of marijuana-related tweets according to gender

<table>
<thead>
<tr>
<th>Marijuana (%)</th>
<th>Person</th>
<th>Number of tweets (%)</th>
<th>Smoking marijuana (%)</th>
<th>Desire to smoke marijuana (%)</th>
<th>Thinking of quitting marijuana (%)</th>
<th>Quitting (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men 1196 (62.5%)</td>
<td>FP</td>
<td>868 (72.6%) of men</td>
<td>677 (78.0)</td>
<td>120 (13.8)</td>
<td>27 (3.1)</td>
<td>44 (5.1)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>328 (27.4%) of men</td>
<td>304 (92.7)</td>
<td>13 (4.0)</td>
<td>8 (2.4)</td>
<td>3 (0.9)</td>
</tr>
<tr>
<td>Women 717 (37.5%)</td>
<td>FP</td>
<td>423 (59%) of women</td>
<td>339 (80.1)</td>
<td>64 (15.1)</td>
<td>8 (1.9)</td>
<td>12 (2.8)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>294 (41%) of women</td>
<td>277 (94.2)</td>
<td>10 (3.4)</td>
<td>7 (2.4)</td>
<td></td>
</tr>
<tr>
<td>Total 1913 (100%)</td>
<td></td>
<td></td>
<td>1597 (83.5)</td>
<td>207 (10.8)</td>
<td>50 (3)</td>
<td>59 (3)</td>
</tr>
</tbody>
</table>

In the data presented in Table 7.4, there is no considerable gender division between men and women in e-cigarette-related tweeting which is similar to the actual e-cigarette prevalence results from studies in Wales among young people (Goldstone et al., 2016; HBSC, 2015). The only noticeable gender difference originates from the number of tweets about ‘quitting e-cigarettes’ which is higher for women. This does not resonate with the literature on e-cigarettes where its use and quitting efforts are similar for both genders (e.g. Vardavas, Filippidis, & Agaku, 2014).

Table 7.4 Descriptive summary of e-cigarette-related tweets according to gender

<table>
<thead>
<tr>
<th>E-cigarettes (%)</th>
<th>Person</th>
<th>Number of tweets (%)</th>
<th>Smoking e-cigarettes (%)</th>
<th>Desire to smoke e-cigarettes (%)</th>
<th>Thinking of quitting e-cigarettes (%)</th>
<th>Quitting e-cigarettes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men 113 (42.5%)</td>
<td>FP</td>
<td>96 (85%) of men</td>
<td>42 (43.8)</td>
<td>5 (5.2)</td>
<td>23 (24.0)</td>
<td>26 (27.1)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>17 (15%) of men</td>
<td>13 (76.5)</td>
<td>2 (11.8)</td>
<td>2 (11.8)</td>
<td></td>
</tr>
<tr>
<td>Women 153 (57.5%)</td>
<td>FP</td>
<td>138 (90%) of women</td>
<td>49 (35.5)</td>
<td>13 (9.4)</td>
<td>34 (24.6)</td>
<td>42 (30.4)</td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>15 (10%) of women</td>
<td>10 (66.7)</td>
<td>4 (26.7)</td>
<td>1 (6.7)</td>
<td></td>
</tr>
<tr>
<td>Total 266 (100%)</td>
<td></td>
<td></td>
<td>114 (42.9)</td>
<td>20 (7.5%)</td>
<td>63 (23.7)</td>
<td>69 (25.9)</td>
</tr>
</tbody>
</table>

The gender division in shisha smoking tweets is reported in Table 7.5. Men tweeted about shisha smoking with others more than women, but the percentages of tweets on smoking action are quite similar. Likewise, the literature on the gender division in shisha smoking is not clear: Jawad & Power (2016) questioned secondary school children in London and found a higher prevalence of occasional shisha smokers among men but, Fuller et al.
(2015) in their Smoking, Drinking and Drug use (SDD) survey among English schoolchildren found no gender difference in shisha smoking prevalence. Yet another study on American adolescents showed a higher prevalence of shisha smoking women (Primack et al., 2015). These studies present different outcomes, and the study presented here merely presents an indication of attention given to shisha on Twitter which signals for space for more research into this smoking habit.

Table 7.5 Descriptive summary of shisha-related tweets according to gender

<table>
<thead>
<tr>
<th>Shisha (%)</th>
<th>Person</th>
<th>Number of tweets (%)</th>
<th>Smoking shisha (%)</th>
<th>Desire to smoke shisha (%)</th>
<th>Thinking of quitting shisha (%)</th>
<th>Quitting shisha (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men 96 (59.3%)</td>
<td>FP</td>
<td>28 (29.2%) of men 68 (70.8%) of men</td>
<td>14 (50.0)</td>
<td>12 (42.9)</td>
<td>2 (7.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>47 (69.1)</td>
<td>19 (27.9)</td>
<td>2 (2.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women 66 (40.7%)</td>
<td>FP</td>
<td>34 (51.5%) of women 32 (48.5%) of women</td>
<td>17 (50.0)</td>
<td>15 (44.1)</td>
<td>2 (5.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not FP</td>
<td>26 (81.2)</td>
<td>6 (18.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total 162 (100%)</td>
<td></td>
<td>104 (64.2)</td>
<td>52 (32.1)</td>
<td>4 (2.5)</td>
<td>2 (1.2)</td>
<td></td>
</tr>
</tbody>
</table>

This section has assessed the smoking tweets content according to gender and results suggest that there are relatively small differences between men and women, but the main difference relates to marijuana smoking. When comparing these findings to the findings of Chapter 6 (section 6.1.2), similar patterns are revealed that whilst women tweet more in general, it is not necessarily about smoking. The results of this chapter and the previous have shown that the gender divide in the tweets of this sample occurs mostly on the level of social media engagement. Linking these results to the gender differences mentioned in the literature review in Chapter 2, there is little evidence to suggest that their experiences of smoking are portrayed distinctively different on Twitter. Overall, the gender variance is too small for policymakers to make a distinction online or provide gender-specific social media intervention.

7.1.3 Smoking content divided by age

Following on from gender, the content of the smoking tweets is analysed according to age and again, is based on the type of smoking product and smoking action. According to Figure 7.1, tobacco-related tweets have the vast majority among all the age groups. The proportion of marijuana-related tweets decreases with age whereas in contrast the posting of e-cigarette-related tweets increases. Shisha has the highest prevalence in the 22-24 age group and is the only age group with more than 2% of tweets on that subject.
Information on smoking action is demonstrated per age group in Figure 7.2. This bar chart illustrates that, apart from the oldest age group, approximately three-quarters of the tweets in each group refer to the action of carrying out smoking. For the oldest age group (i.e. 25 and over) the most substantial proportion of tweets refer to the action of quitting.

Interestingly, the youngest age group do have a very small percentage (10.3%) of tweets referring to quitting, further emphasising that smoking cessation programmes should also target those below 18.

The ‘thinking of quitting’-related tweets have a higher proportion than ‘quitting’-related tweets in the two younger age groups. This difference indicates that the sample members under 22-years-old tweet more content in which they think about quitting smoking but do not tweet about actually starting the attempt. It might be that they never attempt to quit smoking, but it is more likely that they do not tweet about their actual quitting experiences as they, for example, feel they are likely to fail and do not want to be exposed to possible mockery.

On the contrary, the older sample members might be more inclined to tweet about how the quitting attempt is going to receive support from their followers. The proportion of tweets relating to ‘desire to smoke’ is decreasing with age. This is likely due to older sample
members not finding that topic interesting enough to tweet about. Smoking becomes more of a habit than a self-identifying behaviour as people get older (Denscombe, 2001b).

Figure 7.2 Number of smoking type tweets per age group with percentages within each age group in brackets.

This section has revealed that there are age variations in the smoking-related subject matter of the tweets posted by the sample members, both in terms of smoking products and smoking activities. A particularly important finding is that the quitting smoking tweets are present across all age groups; even the individuals below the age of 18 posts about serious quitting smoking attempts.

7.2 Smoking timelines

Attention now moves to investigate the temporal aspects of the smoking-related tweets posted by the sample members. Although tweets have been placed on timelines before (e.g. Bollen, Pepe, & Mao, 2009; Murnane & Counts, 2014), there has been no longer-term temporal analysis of such tweets or any detailed analysis according to smoking products. This depth of temporal analysis is unique to this study and provides a novel insight into the times at which people tweet about smoking and smoking-related ideas and actions. This provides another element to the understanding of the social meaning of smoking so that smoking interventions can be tailored to these young people better.
Research has shown that 70% of current smokers have their first cigarette within one hour of waking (ASH, 2015a), and this may, or may not, prompt an individual to tweet about the activity. There may be other times of the day or parts of the week or year which provide a catalyst for individuals to tweet about smoking. The distribution of tweets on timelines can reveal when there are spikes in the number of tweets, and this can be compared to real-life events. Some literature has triangulated the content analysis with real-life events, e.g. West et al. (2012) demonstrated how most ‘binge drinking’ tweets coincided with alcohol infested occasions such as New Year’s Eve or Halloween. This information can help identify when there is more interest in tweeting smoking-related content and what real-life events are the most likely cause.

Each tweet in the research database is accompanied by a specific time and date (e.g. Tue Jan 13 14:21:53 2015). This information, therefore, allows for a precise analysis of tweeting activity according to hours within days, the days of the week and months of the year. The section presented here outlines when the young people in the sample are tweeting about smoking and what their smoking-related tweets tend to focus on at those specific times.

7.2.1 Comparing smoking timelines to general Twitter activity

When plotting smoking tweets on a timeline, it is difficult to distinguish when a tweet is posted as part of the general Twitter activity or because the content is specific to smoking. Therefore, besides the timelines with smoking tweets, timelines with general Twitter activity are shown. This is done to determine whether there is a different temporal pattern in the use of Twitter for smoking content compared to general tweeting. To feature the difference, a subset of 590,297 tweets were selected from 233 sample members to serve as ‘general Twitter activity’ (see section 5.1.3 of Chapter 5 for the details).

In addition, the smoking tweets considered in this temporal analysis are dated between 1 July 2013 and 30 June 2016. These three years of Twitter feeds were selected as the Twitter gathering program collected tweets from individual users retrospectively until it reached the limit of 3200 meaning that the collected Twitter archives consist of tweets made from June 2016 backwards for each of the sample members. July 2013 was chosen as the lower cut-point as before that time the collected tweets were more sporadic.

Furthermore, the last data collection point was June 2016, and for the monthly timeline, it was important that the first six months of the year were not overrepresented in the results. For the general activity tweets, the cut-off point of three years incorporated 538,341
(91.2%) of all the tweets from the selected 233 sample members which also included 1855 smoking tweets.

The first set of charts, Figures 7.3 to 7.5 illustrate the percentage of tweets relating to each smoking product and the general Twitter activity across hours in the day, days of the week and months of the year, respectively. Notably, these percentages are based on the proportion of each smoking product category to allow direct comparisons across the products and with the general activity.

7.2.1.1 Hourly timelines

The hourly chart (Figure 7.3) shows that most tweets are posted between 9 am and 11 pm. The tobacco-related tweet pattern is comparable to the overall Twitter activity of the sample population except for the evening hours when there is a vast increase in general Twitter activity but not for tobacco-related tweets. Tobacco smoking tweets increase rapidly in the morning and only slowly increase during the day with most tweets posted in the evening hours between 8 pm and 11 pm. The percentage of marijuana-related tweets increases slowly as the day progresses too with occasional spikes which are entirely different from general activity. Marijuana is also the subject in which the percentage has a slower decline in the night hours. The percentage of e-cigarette references distributed over the day has its peak in the morning between 10 am and 1 pm and rapidly declines after 8 pm. Shisha smoking-related tweets have a significant peak in the later hours of the evening.
Figure 7.3 Timeline of the percentage of tweets per hours of the day according to the smoking product referred to in the tweets.

The difference in temporal patterns between the smoking products could be explained by the motives for using the product. Marijuana and shisha smoking are generally part of social activities and, therefore, tweeted about mostly during the ‘social hours’ in the evening. Shisha smoking is an activity to fill up time and to socialise according to a systematic review by Akl et al. (2015) and those times would fit with the results in this sample where tweeting peaks between 6pm and 10 pm. Similar results are found for marijuana smoking which, it is argued, young people prefer to smoke socially (e.g. Tyler, 2015). Similarly, Dunlap et al. (2005) investigated how people smoked marijuana and concluded that the preferred approach is at social occasions which are likely to occur in the evenings. In contrast, tobacco is used for emotional control and helping an individual remain in their ‘normal’ state (J. L. Johnson et al., 2003), and e-cigarettes are used to help quit tobacco smoking (Vardavas et al., 2014). The distribution of tweets associated with tobacco and e-cigarettes could, therefore, relate to times when nicotine dependence is highest which is during the day (8 am to 8 pm when people are awake).

7.2.1.2 Weekly timelines

The frequency of tweets is investigated across the week in Figure 7.3, and there appears to be no particular variability beyond a peak in e-cigarette related tweets on Wednesday. Tobacco and marijuana references are mostly made at the beginning of the week and
slowly decrease towards the weekend, and shisha references are mostly made in the middle of the week. Most interesting is the difference between the distribution of the different smoking-related tweets and those of the general Twitter activity. The prevalence of smoking-related tweets is higher during the week whereas the general Twitter activity is highest during the weekends.

The results of the weekly timeline are perhaps not quite what was expected. For tobacco and e-cigarettes, the prospect was a more equal variance of tweets during the week because those products are more often consumed in isolation than others (Myslín et al., 2013) and based on a craving for nicotine which should be equal on a daily basis. Also, the expectation was that weekends would have more marijuana and shisha referenced tweets as weekends provide more opportunity for individuals to engage in ‘social’ practices that might involve smoking. The graph in Figure 7.4 does not show such a weekend peak. However, it may be that many of the individuals who are smoking marijuana and shisha at weekends choose not to tweet about it then, but possible tweeted about plans to smoke on the weekends during a weekday.

Figure 7.4 Tweet frequency by day of the week according to the smoking product referred to in the tweets.
The patterns shown in Figures 7.3 and 7.4 suggest that individuals tend to post smoking-related tweets (irrespective of product) most often during the weekday evenings. This may or may not reflect a peak in smoking activity itself but would seem to be an optimal time for health organisations to engage with young people via Twitter messages.

7.2.1.3 Monthly timelines

The frequency of smoking tweets per month shown in Figure 7.5 illustrates activity peaks that differ considerably with the monthly pattern of general Twitter activity. This is particularly seen in January and October when there is a marked increase in tobacco and e-cigarette related tweets. These two months are common months within the year for people to stop smoking tobacco: in January, many individuals make New Year’s resolutions to curb a health-damaging behaviour such as overeating or smoking. Likewise, in October there is a national campaign for people to quit smoking at the start of this month (i.e. Stoptober) (Brown et al., 2014; Kotz, Stapleton, Owen, & West, 2011). In the study on young people’s awareness and usage of e-cigarettes in Wales, Goldstone et al. (2016) reported that out of the people that tried an e-cigarette 22.1% used it to quit smoking. Therefore, the peak in e-cigarette related tweets during January and October may be linked to quitting attempts.

The frequency of marijuana tweets is highest in March until June. It is likely that when the smokers gather to smoke marijuana, they tend to do it outside (Dunlap et al., 2005) which is most comfortable in spring and summer. Furthermore, as more people smoke marijuana during these times, it is more visible to others which will likely increase marijuana-related posts too. Shisha smoking peaks in spring and autumn which can have many reasons but most likely, these are the times when people hang out the most and smoke shisha together.
Figure 7.5 Monthly tweeting activity according to the smoking product referenced in the tweets.

The time charts in this section revealed that the smoking-related tweets follow a distinctly different pattern to the overall Twitter activity of the sample with shisha as the overall biggest anomaly. The monthly time series show that most tobacco and e-cigarette related tweets seem to coincide with Stoptober and New Year quit attempts. It would be sensible to increase social media support messaging around these times as young people seem to be optimally engaged in smoking-related social media activity. Marijuana smoking is mostly tweeted about in the warmer months and those times should be targeted for marijuana cessation. Lastly, the shisha time series seem to be rather erratic when compared with the other products and the general Twitter activity. However, it must be noted that the number of tweets relating to shisha smoking is relatively few compared to other products and there may be too few to determine a distinct temporal pattern and hence formulate online cessation support policy.

To investigate whether the time series derived above are related to a more distinct pattern of tobacco smoking prevalence, the following part explores the detail across weeks in the year for tobacco as it is the most common smoking product and the main feature of this study.
7.2.2 Tobacco timelines

Tobacco was the most occurring smoking product in the tweets from the smoking-related sample with 9442 (56.6%) tobacco-related tweets. Furthermore, tobacco is smoked most regularly by young people (ASH Wales, 2010; HBSC, 2015). The time graph (Figure 7.6) illustrates the most noteworthy results showing the distribution of smoking action tweets in the 52 weeks within a year between 1 July 2013 and 30 June 2016 and reveals more information on which periods in the year are most important.

![Figure 7.6 Timeline of the frequency of tobacco-related tweets per week in the years July 2013 to June 2016.](image)

The first noticeable peak in Figure 7.6 is seen in the first week of October. Other increased tobacco-related Twitter activity can be perceived in the first week of the New Year, but this peak is equalled by multiple other upsurges in the first half of the year. Besides the peak in October, tweeting about tobacco is less popular in the second half of the year. This might be due to more people smoking tobacco outside and is, therefore, more visible and more on people’s minds (similar to marijuana mentioned previously). Another option is that even though three whole years were chosen as the data for this time chart that the most tweets that were picked up from the sample members are made in 2016. To get a better notion of what these peaks entail a graph was created with the tobacco monthly timeline based on smoking activity (Figure 7.7).
In Figure 7.7, ‘smoking’ tobacco is most prominent in the spring and summer months when the weather is nicer, and the frequency decreases later in the year. Particularly April has a high percentage of ‘smoking’ and ‘desire to smoke’ related tweets. The graph indicates that, as expected, the ‘thinking of quitting’ content sees an increase around September (pre-Stoptober) and the ‘quitting smoking’ tweets created the peak in October. Interestingly, the number of ‘thinking of quitting’ tweets does not increase in December which seems counter-intuitive as there are many tweets about quitting in January and thinking of quitting tweets do not precede these attempts.

In the previous graphs (Figure 7.6), a spike was visible during the first week of October. As that is the Stoptober month, it is likely that these are related. The largest spike of ‘other’ in May can be explained by a new regulation of standardised packaging for tobacco products which came into effect May 20, 2016. This new regulation requires picture warnings covering 65% of both sides and the top of the package (ASH, 2016b) and affects the young people directly. In October, there is also a small increase of ‘other’ tobacco-related tweets. Here ‘other’ refers to tweets on October 2015 and coincides with the introduction of a new policy to ban individuals from smoking in private vehicles when
transporting children in that vehicle at the same time. This ban affects the sample members less as they are not likely to have children (yet) and it is likely that it was tweeted about less than the picture warning regulation. These results illustrate how Twitter is used to vent about recent events that affect their own lives (Kwak et al., 2010).

In summary, the temporal analyses presented in this section have revealed when smoking-related tweets are posted and particularly the significance of these times. Smoking as a social activity has become evident through the tweets that were made during social hours but do not necessarily relate to general Twitter activity social hours. The case that smoking-related tweets standalone from general Twitter activity can further be seen in the weekly and monthly timelines. Moreover, the tobacco smoking tweets shown over the year illustrate how specific time periods have noteworthy and explainable peaks. These findings increase the understanding of the social meaning of smoking by illustrating when these tweets have enough meaning for individuals to prompt a message to be posted on Twitter.

7.3 Sentiment analysis

Attention now turns towards a deeper understanding of the sentiment contained in the tweets of the study sample. This is achieved by using a widely-used sentiment analysis methodology (see Chapter 5, section 5.2.4) which entails an examination of the words used within a piece of text, in this case, tweets. This study applied the sentiment program Sentistrength which scores each word with a positive and negative number between 1 and 5 for positive sentiments or for negative, -1 to -5 (1 is neutral, and 5 is extreme). The advantage of using such as a methodology is that both sentiments can be present within the same tweet. Most sentiment analysis programs only indicate a direction, i.e. positive or negative.

To uncover the sentiment towards certain topics, an average score is generated for the tweets within each theme. Similar methods were used in a study by Myslín et al. (2013) who used Twitter content to understand smoking behaviour and differing perceptions of emerging tobacco products. Myslín et al. gathered 7362 tweets on different smoking products including tobacco, e-cigarettes, and shisha. Sentiment analysis was undertaken, and the average sentiment score for the tweets was calculated. They concluded from the sentiment analysis that e-cigarettes and shisha are tweeted about more positively than tobacco.
The sentiment measuring program ‘Sentistrength’ (Thelwall et al., 2012) gives the tweet below the sentiment strength of positive 2 and negative -4.

“I hate that I like hookah” (2;-4)

I[0] hate[-3] that[0] I[0] like[1] hookah[0]

This example further shows how a tweet can be both positive (appreciating something) and negative (disliking how much shisha is liked) at the same time which would not have been clear from other sentiment analysis programs that would qualify this tweet as ‘negative’ or, when they are based on emotions, ‘angry’.

A score is given to each of the tweets separately, and for the purpose of the quantitative content analysis presented here, an average is calculated for each category of the smoking product mentioned in the tweets. A single tweet does not reveal much information about smoking and is more likely to refer to the Twitter composition style of the originator. Therefore, average sentiment scores are calculated for each of the smoking products.

### 7.3.1 Sentiment analysis of smoking products

Table 7.6 demonstrates the sentiment analysis of the smoking tweets per product (as mentioned in the tweets) and person (i.e. the first person or not) referred to in the content. The sentiment analysis in this table reveals little variance between the first person smoking and someone else or between the various kinds of smoking action. The sentiment analysis here suggests that, in general, the sample tends to be less positive about the smoking behaviour of others and most positive about e-cigarettes, but this variance is small.

**Table 7.6 Descriptive table on the type of product, who is mentioned in the tweet and the average sentiment score.**

<table>
<thead>
<tr>
<th>Smoking product</th>
<th>Person</th>
<th>Number of tweets</th>
<th>Average positive sentiment</th>
<th>Average negative sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobacco</td>
<td>I</td>
<td>6850</td>
<td>1.47</td>
<td>-1.63</td>
</tr>
<tr>
<td></td>
<td>Not I</td>
<td>1678</td>
<td>1.50</td>
<td>-1.78</td>
</tr>
<tr>
<td><strong>Average for tobacco</strong></td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.65</strong></td>
</tr>
<tr>
<td>Marijuana</td>
<td>I</td>
<td>1291</td>
<td>1.49</td>
<td>-1.84</td>
</tr>
<tr>
<td></td>
<td>Not I</td>
<td>622</td>
<td>1.47</td>
<td>-1.91</td>
</tr>
<tr>
<td><strong>Average for marijuana</strong></td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.87</strong></td>
</tr>
<tr>
<td>E-cigarettes</td>
<td>I</td>
<td>234</td>
<td>1.57</td>
<td>-1.53</td>
</tr>
<tr>
<td></td>
<td>Not I</td>
<td>32</td>
<td>1.56</td>
<td>-1.41</td>
</tr>
<tr>
<td><strong>Average for E-cigs</strong></td>
<td></td>
<td></td>
<td><strong>1.57</strong></td>
<td><strong>-1.54</strong></td>
</tr>
<tr>
<td>Shisha</td>
<td>I</td>
<td>67</td>
<td>1.38</td>
<td>-1.41</td>
</tr>
<tr>
<td></td>
<td>Not I</td>
<td>100</td>
<td>1.54</td>
<td>-1.13</td>
</tr>
<tr>
<td><strong>Average for Shisha</strong></td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.31</strong></td>
</tr>
<tr>
<td>Overall average</td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.62</strong></td>
</tr>
</tbody>
</table>
The average negative sentiment score was more diverse. Tweets concerning marijuana were the most negative with an average score of -1.87 and the tweets on shisha smoking received the least negative score of -1.31 (see Table 7.6). This negativity could have several explanations but most likely mentioning marijuana smoking on Twitter is part of a hip-hop lifestyle that is linked to criminal behaviour (as smoking marijuana is illegal) and rap culture (which often includes swearing) (Kelly, 2005).

In addition to the sentiment analysis according to the person mentioned in the tweet, smoking action across smoking products were computed, and the results are shown in Table 7.7. According to this table, the most positive sentiment score was attached to ‘smoking’ and ‘quitting smoking’ for each type of smoking product. For tobacco, the positive score for quitting was greater than that for smoking which is encouraging as this suggests that young people are more optimistic about quitting tobacco than smoking it.

Table 7.7 Descriptive table on the type of product, the smoking action and the average sentiment score.

<table>
<thead>
<tr>
<th>Smoking Product</th>
<th>Type</th>
<th>Number of tweets</th>
<th>Average positive sentiment</th>
<th>Average negative sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobacco</td>
<td>Smoking</td>
<td>4040</td>
<td>1.48</td>
<td>-1.66</td>
</tr>
<tr>
<td></td>
<td>Desiring to smoke</td>
<td>970</td>
<td>1.34</td>
<td>-1.61</td>
</tr>
<tr>
<td></td>
<td>Thinking of quitting</td>
<td>1192</td>
<td>1.39</td>
<td>-1.66</td>
</tr>
<tr>
<td></td>
<td>Quitting smoking</td>
<td>1848</td>
<td>1.60</td>
<td>-1.54</td>
</tr>
<tr>
<td>Average for tobacco</td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.62</strong></td>
</tr>
<tr>
<td>Marijuana</td>
<td>Smoking</td>
<td>1597</td>
<td>1.49</td>
<td>-1.85</td>
</tr>
<tr>
<td></td>
<td>Desiring to smoke</td>
<td>207</td>
<td>1.37</td>
<td>-1.78</td>
</tr>
<tr>
<td></td>
<td>Thinking of quitting</td>
<td>50</td>
<td>1.46</td>
<td>-1.92</td>
</tr>
<tr>
<td></td>
<td>Quitting smoking</td>
<td>59</td>
<td>1.67</td>
<td>-1.90</td>
</tr>
<tr>
<td>Average for marijuana</td>
<td></td>
<td></td>
<td><strong>1.45</strong></td>
<td><strong>-1.87</strong></td>
</tr>
<tr>
<td>E-cigarettes</td>
<td>Smoking</td>
<td>114</td>
<td>1.66</td>
<td>-1.60</td>
</tr>
<tr>
<td></td>
<td>Desiring to smoke</td>
<td>20</td>
<td>1.30</td>
<td>-1.40</td>
</tr>
<tr>
<td></td>
<td>Thinking of quitting</td>
<td>63</td>
<td>1.49</td>
<td>-1.51</td>
</tr>
<tr>
<td></td>
<td>Quitting smoking</td>
<td>69</td>
<td>1.62</td>
<td>-1.35</td>
</tr>
<tr>
<td>Average for E-cigs</td>
<td></td>
<td></td>
<td><strong>1.57</strong></td>
<td><strong>-1.54</strong></td>
</tr>
<tr>
<td>Shisha</td>
<td>Smoking</td>
<td>104</td>
<td>1.50</td>
<td>-1.21</td>
</tr>
<tr>
<td></td>
<td>Desiring to smoke</td>
<td>52</td>
<td>1.37</td>
<td>-1.21</td>
</tr>
<tr>
<td></td>
<td>Thinking of quitting</td>
<td>4</td>
<td>1.50</td>
<td>-1.50</td>
</tr>
<tr>
<td></td>
<td>Quitting smoking</td>
<td>2</td>
<td>1.50</td>
<td>-1.00</td>
</tr>
<tr>
<td>Average for Shisha</td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.31</strong></td>
</tr>
<tr>
<td>Overall average</td>
<td></td>
<td></td>
<td><strong>1.47</strong></td>
<td><strong>-1.62</strong></td>
</tr>
</tbody>
</table>

The most negative sentiment scores in Table 7.7 are given to marijuana-related tweets. Even the smoking of marijuana is found to be more negative than the smoking of any other product. However, this might be related to the association of marijuana with a certain
lifestyle explained earlier which makes the tweets come out more negative in the sentiment analysis then intended. When reading through the most negatively scored marijuana-related tweets, the negative sentiment might not be meant negative at all but simply include negatively associated pop-culture references;

“I totally agree with all you weed smoking dreadlock motherf*ckers about no wars and all that unrealistic stuff but it won't happen.” (1;-5)

The ‘Sentistrength’ program added (as shown in the example above) a ‘-3’ to the word ‘dreadlock’. There are more words that are not meant negatively but are classified as such by the program, and these words are more commonly used in sentences on the topic of marijuana than another smoking product. These negatively associated words skew the results. This last notion was also found in the Twitter analysis of marijuana content by Cavazos-Rehg et al. (2015) who used a crowdsourcing version of sentiment analysis program. They discussed how even though they were filled with negative words, 77% of the marijuana-related tweets had positive content. As shown by this example of Cavazos-Rehg et al., marijuana Twitter content is a particularly difficult case for sentiment analysis and the results should be considered with that limitation.

In Table 7.8, there is only a small difference between men and women in their sentiment scores towards types of smoking products. The only interesting outcome is that women score higher on all sentiments calculations except the positive score of shisha. These results indicate that women tweet about smoking with more sentiment (both positive and negative) than men. The largest variance between the average sentiment for men and women is present in the category of marijuana smoking. Relating this to the results in section 7.1.2 earlier in this chapter, marijuana smoking provides the only gender difference, but it is small. Women post fewer tweets with marijuana content but when they do it is more often about others smoking marijuana, and the sentiment score is overall more negative.
The average sentiment scores across smoking products demonstrate how quitting is associated with more positive sentiment than smoking. Moreover, people are most negative about marijuana smoking. Attention should be placed on the e-cigarettes and shisha as they are associated with both high positive and low negative sentiment scores which have also been found in Myslíň et al.’s (2013) article where the authors discussed how the lack of regulations on e-cigarettes and shisha puts those products in a more positive light than tobacco. The number of tweets in this sample and the usage of these products are relatively low (Fuller et al., 2014; HBSC, 2015) but this is likely to increase, and future work should focus in more detail on these products and consider the implications for smoking cessation intervention and policy.

### 7.4 Temporal analyses of sentiment scores

The last part of this chapter reports on the results from temporal analysis of the sentiment scores presented above and considers their variation across hours within the day, days within the week and months of the year. For this analysis, the absolute value of the negative sentiment scores is used (i.e. the negative sign is ignored). This allows for a more intuitive display of the peaks and troughs within the time series graphs of sentiment scores compared to the general Twitter activity sentiment scores. Also, the smoking products are split up; there are separate graphs for tobacco and e-cigarettes and marijuana and shisha.

The time series chart (Figures 7.8) demonstrates the sentiment scores during the hours of the day calculated for tobacco and e-cigarette tweets. Figure 7.9 reveals the sentiment score on an hourly timeline for marijuana and shisha-related tweets. Both time series charts illustrate that the general positive and negative sentiment are almost equal during any part of the day, but the sample members tweet more negatively than positively (except for shisha-related tweet after 2 pm). The sentiment analysis reveals that most fluctuation occurs in the night hours, and in those hours when fewer tweets are made, the negative sentiment scores are more extreme.
The variability of average sentiment across the 24 hour period lessens when the average sentiment is shown for days across the week as is illustrated in Figure 7.10 and 7.11, suggesting that there is more variability within the day than across the week. The negative sentiment scores are higher than the positive sentiment scores in the smoking-related tweets throughout the week, except for shisha and occasionally e-cigarettes. The sentiment scores for tobacco do not alter much during the week. E-cigarette tweet sentiment fluctuates during the week, but young people tweet most positively about it on Monday.
and Thursday, and most negatively during the weekend (Friday, Saturday, and Sunday).
Marijuana tweets have the strongest negative sentiment throughout the week with increases on Tuesday, Wednesday and Saturday. In general, shisha smoking has the lowest negative and positive sentiment. These observations prompt that the sample members express more emotion in their smoking tweets during the weekend when there are fewer tweets made (see Figure 7.3 earlier).

**Figure 7.10** Time series of average sentiment score for tobacco and e-cigarette tweets per days of the week

**Figure 7.11** Time series of average sentiment score for marijuana and shisha tweets per days of the week

The last figures (7.12 and 7.13) of this chapter report the average sentiment of tweets divided by smoking type per month of the year. Positive sentiment about e-cigarettes picks
up around September and October which coincides with the Stoptober movement. Tobacco smoking tweets show, similar to the previous temporal analysis, no remarkable fluctuation throughout the year and the positive scores follow almost identical patterns to that of the negative scores for e-cigarette related tweets. Comparing the positive and negative scores of e-cigarette-related tweets reveals that over the year the sample members tweet more positively about it than negatively. In Figure, 7.13, there is little change in sentiment scores in the first months of the year, but from April the negative tweets about marijuana start to decrease, and the tweets have a higher positive score than negative. Shisha referenced tweets are most positive throughout the year stands alone above all the other sentiments scores at each month.

Figure 7.12. Timeline of average sentiment score for the tobacco and e-cigarette tweets per months of the year

Figure 7.13. Timeline of average sentiment score for the marijuana and shisha tweets per months of the year.
7.5 Concluding remarks

The first objective of this chapter was to examine, in more detail, the different types of engagement with the Twitter feed, including differences in the subject content of tweets posted. This was accomplished by quantifying the extent to which different smoking products were mentioned, whether the tweet concerned the first person or some other subject and some assessment of the smoking-related activity. This detailed, empirical scrutiny of smoking-related tweets has not been previously undertaken.

The results reveal that most of the sample members tweet about smoking tobacco which is as expected given the dominance of this type of smoking product compared to the other smoking-related products. This observation is also supported by wider literature (ASH, 2015a; Gilmore et al., 2015; Myslín et al., 2013).

Assumed gender and age were added to the quantitative content analysis. As in the previous chapter, there was no clear division between men and women and smoking-related content. However, the age analysis presented an important finding in terms of policy relevance for The Filter Wales program as there were tweets about quitting smoking from the youngest age group (i.e. 18 years old and below). This suggests that thoughts about quitting are manifest across all young people, not just the older groups.

The second objective was to establish the temporal and sentimental qualities of the smoking-related tweets. The temporal analyses indicated that tobacco and e-cigarettes are tweeted about at random times during the day, and most tweets about tobacco and e-cigarettes occur in January and October. This last finding suggests that the sample members participate in, or at least are aware of, the anti-smoking campaigns which dominate at these times of the year, thus prompting individuals to tweet about quitting tobacco or e-cigarette smoking during these periods. Marijuana and shisha smoking tweets were most popular during ‘social’ hours which indicate that these smoking products are part of social activities which has also been witnessed in other research (e.g. Akl et al., 2013; Dunlap, Johnson, Benoit, & Sifaneck, 2005; Jawad, McIver, & Iqbal, 2014).

The smoking product specific time series were visually compared to the general Twitter activity across the 233 smoking sample members. The results show that the smoking-related timelines are considerably different from those of general Twitter activity and it can be argued that there is particular importance in posting smoking-related content at certain times. These finding could be used to select a time to implement a social network based
anti-tobacco campaign which would be most effective in the evening hours and the beginning of spring (in addition to the already existing anti-tobacco campaigns).

The longer duration time series analysis during all weeks in a year and the time series on tobacco smoking activity (from July 2013 through to June 2016) revealed some more distinct points of interest. It revealed that, besides tweeting about their own experiences regarding tobacco smoking, sample members also post content about tobacco regulation. This implies that awareness of tobacco policy and control is relatively high, and it may be mentioned on Twitter if it concerns their everyday experiences.

The second part of the second objective concerned the sentimental qualities of tweet content. The sentiment analyses established that young people are, in the main, not positive in their posts about smoking, but they are more positive when it comes to quitting. Overall, the tweets about marijuana smoking received the most negative sentiment, but this result might be more related to the association of marijuana with criminal behaviour and rap culture. The sentiment scores placed on timelines showed that tweeting with negative sentiment was overall more common.

Most alarmingly, the tweets about shisha smoking received the most positive and least negative score indicating that shisha smoking is associated with favourable attitudes. This positive sentiment towards shisha smoking also comes through in the sentiment score timelines. However, the total number of tweets about shisha are relatively small and so caution must be employed when deriving conclusions concerning this smoking product.

The following two chapters demonstrate how the Twitter data of the sample members can be scrutinised by qualitative methods of analysis to provide a qualitative understanding of the social meaning of smoking that originates in the tweets. The next result chapter (8) continues with the content analysis of the tweets and views the smoking-related tweets in combination with other health behaviours within the same tweet. This allows for an in-depth qualitative insight into the possible motives behind the tweets. The final results chapter is Chapter 9 in which the Twitter context of the smoking-related tweets is discussed on the basis of a co-behavioural subsample.
Chapter 8. Comparing health behaviours

This chapter provides an account of the interplay between smoking and other health risk behaviours discussed as the ‘holy four’ by Martin & McQueen (1989), i.e. smoking, alcohol use, healthy eating, and physical activity. This chapter aims to uncover how smoking is connected to other health risk behaviours through an analysis of the tweet content generated by the youth sample. The objective that will be discussed explicitly in this section of the results is:

- to evaluate the extent to which co-behaviours are present in the smoking-related tweets.

The results of this chapter are an exploration of the interplay between smoking and the other health behaviours as described in Chapter 2 section 2.6 within the tweets of this sample. According to Peter et al. (2009) and Wiefferink et al. (2006), health co-behaviour can be considered from two different perspectives: first, engaging in unhealthy co-behaviours for a ‘quick fix’ (e.g. smoking, drinking alcohol, and eating unhealthy food to receive an immediate positive effect on mental well-being) and second, undertaking healthy co-behaviours (e.g. quitting smoking and/or alcohol, dieting, and engaging in physical exercise) to feel more physically healthy. Furthermore, Chapter 2 has expanded on the motives for smoking that are shared with other health behaviours such as engaging in smoking and consuming alcohol while socialising with friends and initiating smoking to lose weight. In this chapter, the combinations and interplay of these health co-behaviours are discussed to further uncover the social meaning of smoking by relating it to other health behaviours.

Another purpose of this chapter is to establish which health behaviours form clusters of activity and in what ways. Scarborough et al. (2011) calculated the economic burden of the illnesses related to the four health risk behaviours (e.g. illnesses such lung cancer, diabetes, and cardiovascular diseases) in the UK in 2006/07. They found that the total economic burden came down to £81.3 billion in 2006/07 from health issues caused by poor health behaviours (Scarborough et al., 2011). It is thought that if health behaviours are part of a combined targeting campaign, the campaign will become more effective, and the overall economic burden will be reduced significantly (Wiefferink et al., 2006). Therefore, this chapter is concerned with the combinations of health co-behaviours mentioned in tweets to better understand how they are united and how targeted interventions could be more effective.
The first section (8.1) is an exploratory overview of the presence of co-behaviours reviewed through a gender and age lens to reveal if there are variations in how different groups connect smoking to one or more health risk behaviours. Moreover, this section presents the sentiment analysis of the smoking co-behavioural tweets to show which combinations present the most positive or negative average sentiment. This is followed by in-depth examinations of the connections in the different type of co-behaviour (sections 8.2 to 8.4).

8.1 Overview of co-behaviour content

The tweets used in this section of the results consist of smoking tweets that have a reference to another health risk behaviour in their content. The co-behavioural subsamples were created through a search for specific words within the smoking-related tweet table indicating the particular health risk behaviour, e.g. each tweet containing the word ‘vodka’ was placed in the subsample for smoking and alcohol consumption. All the search words are described in Chapter 4, section 4.5.4 along with a full description of the way in which the search terms were utilised.

The co-behaviour tweets were placed into specific clusters based on their health behaviour. The methods used in the qualitative content analysis are described in Chapter 5.2.2. In essence, this section of the results employed an inductive process in which the co-behaviour tweets were separated based on health behaviour and coded into themes. The results suggested 11 themes for alcohol co-behaviour, 11 for healthy eating co-behaviour, 6 for physical exercise co-behaviour, and 5 multiple health co-behaviours (all themes are represented in the tables for each health co-behaviour in Appendix D). These themes were further reduced to three core themes per co-behaviour to keep the content analysis manageable. These three core themes for alcohol co-behaviour (which are described in section 8.2) are the connection between alcohol and smoking, the social settings in which they are consumed, and the purchase of the two products. Healthy eating and physical exercise are combined in the content analysis as the physical exercise co-behaviour had only a small number of tweets to discuss in isolation, and both had the same three core themes (outlined in section 8.3) which are smoking and unhealthy co-behaviour, quitting smoking and healthy co-behaviours, and the combination of one healthy and one unhealthy behaviour. The three themes for the multiple co-behaviours (which are represented in section 8.4) are combining unhealthy behaviours, combining healthy behaviours, and comparing the health behaviours to each other.
8.1.1 The occurrence of co-behaviours

The vast majority of the smoking-related tweets contained no reference to another health behaviour (15410 tweets, 92%). From the 8% that did, 1235 (7.4%) had a reference to one other health behaviour in combination with smoking, while the remaining 0.6% (105 tweets) had two or more co-behaviours present. As is apparent from Figure 8.1, tweets containing content on alcohol co-behaviour occurred most frequently, followed by tweets containing text about healthy eating co-behaviour. In the category of multiple health co-behaviours, smoking, alcohol and healthy eating was the most frequent type of tweet.

![Figure 8.1 Prevalence of other health behaviour references in smoking-related tweets]

8.1.2 Co-behaviour and gender

The distribution of co-behavioural tweets differed between men and women which is illustrated in Table 8.1. The total number of tweets with smoking and (at least one) other health behaviour reference is 1340, with 658 (49.1%) of these tweets belonging to men and 682 (50.9%) tweets posted by women. Table 8.1 reveals that men have a higher number of tweets in the alcohol co-behaviour tweets, but women have a higher absolute count in all the other classifications. The proportions of tweets for each health risk co-behaviour are relatively equal.
According to health behaviour literature, women tend to focus more on healthy eating while men think more often of physical fitness (Amos & Bostock, 2007; Mistry et al., 2009; Steptoe et al., 2002). Moreover, women who want to change their lifestyle would likely focus on getting more physical activity as most of them exercise below guidelines (Amos & Bostock, 2007). It is likely that the increased focus is translated into a higher number of tweet containing physical exercise and healthy eating for women in this sample.

Table 8.1 Prevalence of health behaviours in smoking-related tweets divided by gender. Row percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Gender</th>
<th>Alcohol reference (%)</th>
<th>Healthy eating reference (%)</th>
<th>Physical exercise reference (%)</th>
<th>Alcohol and healthy eating reference (%)</th>
<th>Alcohol and physical exercise reference (%)</th>
<th>Healthy eating and physical exercising reference (%)</th>
<th>Alcohol, healthy eating and physical exercise reference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men 658 (49.1%)</td>
<td>373 (56.7)</td>
<td>173 (26.3)</td>
<td>64 (9.7)</td>
<td>26 (4)</td>
<td>6 (0.9)</td>
<td>15 (2.3)</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Women 682 (50.9%)</td>
<td>336 (49.3)</td>
<td>206 (30.2)</td>
<td>83 (12.2)</td>
<td>31 (4.5)</td>
<td>5 (0.7)</td>
<td>17 (2.5)</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>Total 1340 (100%)</td>
<td>709 (52.9)</td>
<td>379 (28.3)</td>
<td>147 (11)</td>
<td>57 (4.3)</td>
<td>11 (0.8)</td>
<td>32 (2.4)</td>
<td>5 (0.4)</td>
</tr>
</tbody>
</table>

A study by Mistry et al. (2009) among Californian youth found similar results; ‘risk takers’ (those who have multiple unhealthy behaviours including drinking alcohol and binge eating) comprise of only 7% of the men and 15.9% of the women. However, the number of smokers in the ‘risk takers’ category was twice as high for men (60%) compared to women (26.9%). Their results indicate that women cluster alcohol use, unhealthy eating, and physical inactivity but not necessarily smoking whereas men have a higher clustering rate of alcohol and smoking (Mistry et al., 2009). Overall, women are more likely to participate in unhealthy co-behaviours than men, but these do not necessarily relate to smoking.

The women in the sample tweeted three times more than men about unhealthy behaviour after they quit smoking whereas men’s tweets contained more references to the consequences of smoking on health as these examples show:

“So i've now quit smoking for 25 days however food intake has increased dramatically #gonnagetfat” (posted by a woman)
“*goes for a run* Omg I should really stop smoking I would run so much better
*gets in from run and has a cigarette*” (posted by a man)

Noteworthy, when analysing the content of the tweets posted by men and women, women tended to provide context to their actions. For example, women mentioned quitting smoking and drinking alcohol or dieting and getting more physical exercise as a New Year’s, dry January or as a Lent resolution while men just declared their intention to quit. As a comparison, two of the women and two of the men in the sample tweeted about their quitting attempt:

“21 days without a drop of alcohol or a cigarette !!! This dry January is easy !!!”
(posted by a woman)

“New year I’m deffo going on a health kick no alcohol no fags and exercise Gunna be healthy for my men” (posted by a woman)

“6 days of no smoking or drinking. Still feel pretty shit but it’s gotta be worth it, right? #Detox.” (posted by a man)

“no smoking no drinking no partying no nothing for a while. I need to sort myself out. Fancy a diet too.” (posted by a man)

This difference in clustering of behaviours can have implications for school interventions as the emphasis for women on a healthy lifestyle is different from that of men who focus more on the combination of smoking and alcohol. However, this content variance can also be related to the way men and women tweet and what kind of tweets get the most desired response from their followers (Kwak et al., 2010).

8.1.3 Co-behaviour and age

The results of the content analysis of co-behaviour by age group are illustrated in Table 8.2. As mentioned in Chapter 4 section 4.4.4, only 27.4% of the people in the sample had included a reference to their age in their Twitter profile. There were 1340 co-behavioural tweets in the sample and of these tweets, 379 (28%) had an age reference. The low proportion of tweets suitable for this analysis indicates that the following results are not necessarily a good representation of the age groups but merely an indication of what are the most common health behavioural combinations of the older and younger members in this sample.
Table 8.2 Prevalence of health behaviours in smoking-related tweets divided by age group. Row percentages are given between brackets (%).

<table>
<thead>
<tr>
<th>Age group</th>
<th>Alcohol reference (%)</th>
<th>Healthy eating reference (%)</th>
<th>Physical exercise reference (%)</th>
<th>Alcohol and healthy eating reference (%)</th>
<th>Alcohol and physical exercise reference (%)</th>
<th>Healthy eating and physical exercising reference (%)</th>
<th>Alcohol, healthy eating and physical exercise reference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 and under</td>
<td>17 (63)</td>
<td>3 (11.1)</td>
<td>5 (18.5)</td>
<td>1 (3.7)</td>
<td>1 (3.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 (7.1%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-21</td>
<td>84 (64.1)</td>
<td>33 (25.2)</td>
<td>9 (6.9)</td>
<td>4 (2.5)</td>
<td>1 (0.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>131 (34.6%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-24</td>
<td>82 (50.6)</td>
<td>47 (29)</td>
<td>22 (13.6)</td>
<td>4 (2.5)</td>
<td>4 (2.5)</td>
<td></td>
<td>3 (1.9)</td>
</tr>
<tr>
<td>162 (42.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 and over</td>
<td>25 (42.4)</td>
<td>18 (30.5)</td>
<td>5 (8.5)</td>
<td>4 (6.8)</td>
<td>1 (1.7)</td>
<td></td>
<td>6 (10.2)</td>
</tr>
<tr>
<td>59 (15.6%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>208 (54.9)</td>
<td>101 (26.6)</td>
<td>41 (10.8)</td>
<td>13 (3.4)</td>
<td>3 (0.8)</td>
<td>10 (2.6)</td>
<td>3 (0.8)</td>
</tr>
<tr>
<td>379 (100%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8.2 illustrates that the highest proportion of smoking tweets with an alcohol reference are made by the two youngest age groups. In the healthy eating and physical activity classifications, the older age groups have a higher proportion of the tweets. With age, tweeting about a healthy lifestyle becomes more popular, and specifically, tweets that have more than one health co-behaviour.

When analysing the posts from the youngest age group in this sample, they seemed mostly impulsive and focused on the young person themselves as was also argued by Livingstone (2002) in her study on self-expression on social network sites. Specific to this age group is that they are considered minors and not allowed to buy cigarettes or alcohol legally and this is reflected in some of the tweets:

“gasping for a fag but my brother's gone out drinking :’(“(17-year-old)

As soon as smoking and alcohol consumption becomes legal, the tweets relate more to going out with friends as the tweet below indicates:

“Anyone fancy going to a beer garden #needstobedone #beer #fags #sun.” (21-year-old)
Moreover, as the sample members get older, the tweets are more based on updating their followers. The tweets from the two older age groups are more informative and outward looking than those from the younger sample members:

“So university turned me to smoking now I’ve quit I really fancy a pint please don’t turn me into an alcoholic!” (25-year-old)

“RT @#####: Lack of exercise kills just as many people as smoking.” (22-year-old)

_I’m obviously an addict; fags, chocolate, food, t.v. why can’t I get addicted to lettuce and exercise?_ (24-year-old)

There are some differences between age groups such as expressing emotions about co-behaviour versus updates of their own behaviour. This notion relates to the motives to use Twitter from other literature as the younger Twitter users tend to use it to connect with others whereas as people get older, their behaviour changes towards presenting more informative content (D. J. Hughes et al., 2012; P. R. Johnson & Yang, 2009).

### 8.1.4 Co-behaviour and sentiment analysis

Sentiment analysis scores pieces of texts and grades each word with a positive and negative number and has been used and explained in section 5.2.4 of Chapter 5 and section 7.3 of Chapter 7. The scores of the analysis are created to outline an overall positive and negative sentiment score as both can occur in the same sentence. The minimum score is 1 and -1 when there are no particular words with sentiment in the text. The next example shows the very positive tweets with a sentiment score of 5 due to word combination ‘f*cking awesome’ and has no negative words and therefore a standard -1 negative score:

_“3 stone down and 3 months of no smoking.. i feel f*cking awesome
#nostoppinghere” (5: -1)_

3[0] stone[0] down[0] and[0] 3[0] months[0] of[0] no[0] smoking[0]. i[0] feel[0] f*cking[0] awesome[2] [2 LastWordBoosterStrength]#nostoppinghere[0]

Each tweet is scored individually and calculating the average of the grouped tweets presents an overall perception of what the young people in the sample feel about each particular type of co-behaviour. The sentiment of each tweet can give an interesting result but is highly individual to the originator. With the intention of providing a view of the sentiment across the interplay of the different co-behaviours, it is important to show the
average of the different tweets. Using Sentistrength (Thelwall et al., 2012) for the qualitative analysis has the limitation that it only provides positive and negative scores rather than a more in-depth analysis of the emotion or sentiment in these co-behavioural tweets. The sentiment scores given are, therefore, scrutinised manually to uncover which element of the tweet (e.g. alcohol co-use or quitting smoking) caused this particular score.

The results of this sentiment analysis of these co-behaviour tweets contain similarities to the sentiment scores in Chapter 7, section 7.3; quitting smoking as well as becoming healthier overall was tweeted about with the most positive sentiment. The content reveals that the sample population tweet uplifting sentiments about their achievements. However, the average sentiment scores also expose how quitting smoking is accompanied by both high positive and negative sentiment. Twitter presents a great platform to show off personal success but also a great platform to express irritation (Hughes, Rowe, Batey, & Lee, 2012; Marwick & Mand, 2010). The originator of the following example is pleased that quitting smoking is going well but also reveals an element of struggle which increased the negative sentiment score:

“’Tis an extremely lucky thing that Mr Jones is away whilst I attempt to give up smoking and sugar and wine for Lent. I'm not nice this week.” (4, -2)

The negative sentiment scored highest amongst tweets in which smoking and other health behaviours are compared. The tweet below illustrates the perceived unfairness young people in the sample seem to experience:

“Fifteen thousand people die from alcohol-related diseases every year in the UK alone. Nobody has ever died from smoking weed.” (1, -3)

The sentiment scores reveal how the average sentiment is dependent on the topic in the tweets and therefore a further exploration into these tweets will reveal more about how these health co-behaviours interplay. The next sections provide in-depth views on the content of the tweets that accompany each co-behaviour.

8.2 Smoking and alcohol consumption: an in-depth focus

As seen previously in Figure 8.1, alcohol consumption is the most commonly mentioned co-behaviour. Other studies have also found a strong link between the use of alcohol and other drugs and smoking among young people (e.g. Bancej, O’Loughlin, Platt, Paradis, & Gervais, 2007; C. C. Johnson, Webber, Myers, Boris, & Berenson, 2009; Leatherdale & Ahmed, 2010). In the Smoking, Drinking, and Drugs lifestyle questionnaire study of
English 15-year-olds in 2015, 90% of the young people who said they smoked in the last week also reported that they drank alcohol or took drugs and the people that said they drank in the last week were 19 times more likely to also be regular smokers (Fuller et al., 2014). In this way, consuming alcohol interlinks to smoking and the following tweets illustrate that too:

“bottle of wine, cigarettes and trashy Netflix shows is my life”

“always need a fag with a drink”

8.2.1 How smoking and alcohol use are connected

There are no clear directions of the pathways between youth smoking and alcohol use as evidence suggests from the literature that they both influence each other. For example, previous studies have found that alcohol consumption is linked to the initiation of smoking regardless of age group or gender (Chen et al., 2002; Jackson et al., 2002). Likewise, the content of the tweets in this sample gave no indication of pathways for initiation of smoking and alcohol, just the combination of the two such as exemplified here:

“I remember when I used to take sport so seriously in high school, then alcohol, smoking and girls came into my life...”

The content of the co-behaviour tweets provides insights into how these two behaviours are often undertaken for similar reasons. They are both used for relaxation, reducing stress, and general happiness: an observation that has also been made by Little (2000). When people want an instant fix for their emotional state, they are likely to engage in ‘quick fix’ health risk behaviours such as alcohol use and smoking (Peters et al., 2009). This co-behaviour theme was the largest in the sample with 389 tweets on this topic. Here are some examples of tweet content which illustrate these perceptions:

“Fag and an alcoholic drink I think! #relaxing” (relaxation)

“I need a double gin/ triple gin/ a bottle of gin and a cigarette. ASAP” (anti-stress)

“living the dream at the moment, sun’s out, BBQ’s out, beers out, cigarette’s out, flip flops out #brill” (general happiness)

Another theme in the smoking and alcohol tweets was quitting both tobacco and alcohol with 85 tweets on this topic. Both behaviours have adverse effects on young people's lives and quitting them both seems beneficial. These tweets present the feelings of remorse experienced after smoking and drinking:
“Never going to drink a zwack bomb or smoke a cig ever again #euuurreggghhh”

“Why do I always think smoking is a good idea when I drink?
#grimmorningbreath”

Unfortunately, these feelings were often forgotten in time, and the benefits mentioned by Little (2000) earlier concerning a ‘quick emotional fix’ will become important again.

Once the person is engaged in an attempt to quit smoking and alcohol, the tweets have a different tone, either positive or negative. People express their negativity increasing with quitting as it causes unwanted feelings and the whole attempt could fail because of it:

“GS has it right, stopping drinking and smoking won't make me live longer, it'll just seem longer”

“It's been almost week now and not a drop of alcohol or a single drag of a cigarette. I haven't murdered anyone yet! #watchout #stoptober”

However, as Twitter is a platform on which people give updates on their success (Zhao & Rosson, 2009), the positive effects of quitting both is tweeted about more often than the negative effects. These sample members have a positive outlook on the quitting smoking and alcohol:

“Turns out quitting smoking, marijuana and drinking alcohol for a week has made me the happiest and most self-confident I have been in years.”

“Can't believe it's been 16 months since I last had a cigarette or an alcoholic drink! Best choice I ever made! :D #feelinggood #willpower”

In several tweets, young people tweet about having an element of their lives change (e.g. new job, break-up, or finishing education) and they chose to change their smoking and alcohol behaviour at the same time:

“@############ I actually don't drink much since I've worked at miss millies tbf to myself plus ive gone a week without a cigarette go B !! X”.

“Just gonna get my head down in college, go to uni hopefully, get the job I want, need to stop smoking and drinking big time.”

Quitting smoking and reducing alcohol intake are occasionally accompanied by other changes to lifestyle such as dieting and increasing exercise which will be explained in more detail in section 8.4.2 of this chapter.
Alcohol has the added inconvenience that it can cloud a person’s judgement and people give in to things easier than they would if they had been sober (Little, 2000). In this sense, alcohol ‘interferes’ with the quitting smoking attempt, leading to a number of adults in Wales declaring in the lifestyle survey that they have quit smoking but relapse when they are drinking alcohol (Parry et al., 2010). There are 28 tweets in this sample that refer to the challenges of attempting to quit one behaviour but not the other. These tweets illustrate how young people ‘give in’ to the temptation of having a smoke because they also have an alcoholic drink:

“no way! I quit smoking except for when I'm drinking”

“I've relapsed. It's 11 am and I have a glass of wine in one hand and a cigarette in the other #happysunday”.

8.2.2 Social situations with smoking and alcohol

The complex connection between smoking and drinking is also illustrated regarding the social contexts in which they are both consumed. Both behaviours are often undertaken during various social activities such as parties and hanging out at home with friends, and they are consumed together for the group to entertain themselves (Laurier, McKie, & Goodwin, 2000; Little, 2000; Nichter et al., 2010). Several reasons for smoking and alcohol co-use are making social interactions easier as it ‘fits’ with the party scene (Nichter et al., 2010). In the 796 tweets from the sample about smoking and alcohol use, 54 tweets contained a reference to the difficulty of not having a cigarette while enjoying alcohol in these situations. The key influencers here are friends. Friends (both online and offline) are an essential element in the likelihood of someone smoking and drinking alcohol in any social situation (Huang et al., 2014). The sample members like to drink when they go out but having quit smoking earlier puts this person's perseverance to the test:

“I'm trying to quit smoking, but I worry this strong craving I get whenever I drink alcohol will never fade. #quitsmoking #smoking”

“First night of going out, having a drink and not smoking! Gonna be a tough one, thank god the bf doesn't smoke #entertainme”

8.2.3 Expenses and buying of smoking products and alcohol

This subsection of the smoking and alcohol tweets relates not to the consumption, but the purchasing of tobacco and alcohol products. The government taxes both alcohol and cigarettes highly, and the high taxation seems to have a small but significant effect on the
desire to quit smoking (Rice, Godfrey, & Slack, 2010). This desire was also represented in
the tweets of the young people in the sample. The 34 tweets under this theme relate to the
downside of buying alcohol and smoking products and how all the money is being spent on
those products:

“Sick of spending all my money on alcohol and cigarettes not gonna lie”.

“I always spend my loan on tattoos, weed and alcohol.”

The literature on the tax increase of tobacco suggests that young people smoke less and are
less likely to initiate smoking (Lantz et al., 2000; Rice et al., 2010; Van Hasselt et al.,
2015). However, the tweets in this sample indicate that it only affects what they can spend
their money on and the revelation of what a waste of money it is:

“I’ve spent £60 since Christmas to now, mainly on food, alcohol and fags #help”

“If I saved up all the money that I’ve spent on food fags and alcohol I’d be rich
right now”

An element of buying cigarettes and alcohol that became apparent in the literature review
of Chapter 3 was the proxy method in which an adult was asked to purchase the products
for minors (ASH Wales, 2010; Fuller et al., 2014; Robinson & Amos, 2010). On Twitter,
this is portrayed as more of a lazy request from the Twitter users to their followers:

“Can’t muster the energy to go to the shop for a bottle of wine and cigarettes.”

The co-behaviour of smoking and alcohol consumption had the most substantial number of
tweets in the sample, and by far most of the content relates to consuming both at the same
time. The individuals in the sample make plans to quit one or both, but this commonly
remains an idea that can get discarded quickly in social events or when there is a need for
instant emotional regulating. Moreover, the high taxation of cigarettes and alcohol
inconveniences individuals, rather than restricting consumption to a large extent.

8.3 Smoking and healthy co-behaviours: an in-depth focus

After the content analysis of the tweets in the ‘healthy eating’ and ‘physical exercise’
categories, it became clear that their origins and connections are similar and could be
assessed in the same three core themes. Moreover, the number of tweets concerning these
coo-behaviours are (relative to alcohol) quite small and combining them into one type of co-
behaviour was beneficial for the analysis. Therefore, the 496 tweets with healthy eating
and smoking co-behaviour reference and the 159 tweets from the physical exercise and
smoking co-behaviour table are combined for the in-depth focus. Mistry et al. (2009) studied clusters of unhealthy co-behaviours in adolescents in California and found that high physical activity and healthy eating were clustered to illustrate healthy co-behaviour. Therefore, in the following section, the physical exercise and healthy eating co-behaviours are categorised together as ‘healthy co-behaviour’ and consequently low physical exercise and unhealthy eating are considered ‘unhealthy co-behaviour’.

8.3.1 Smoking and unhealthy co-behaviour

The 183 tweets with both a smoking and unhealthy lifestyle reference illustrate that having the two together is most common. Similar to alcohol consumption, unhealthy eating provides a ‘quick fix’ and relaxing experience in combination with smoking (Peters et al., 2009). The shared theme relates to the comfort of no physical activity or dieting:

“A cigarette after a fat ass meal is the best cigarette”

Likewise, Van Lenthe et al. (2009) discussed how perceived stress increases the chance of unhealthy behaviours. Not only is there a desire for smoking and unhealthy eating, but the unhealthy co-behaviour impulse also becomes greater when people quit smoking. The habits related to smoking need to be replaced with other habits, and eating has become the emotion regulator that replaced smoking. The 79 tweets on this topic in the cluster are represented by the following examples:

“Day 3 of no cigs - I'm eating everything in sight”

“since quitting smoking I have to take chocolate with me everytime i go outside... step aside lung cancer and make way for heart disease”

In contrast to other studies which highlight the association of smoking with weight loss (e.g. Allbutt, Amos, & Cunningham-Burley, 1995; Amos & Bostock, 2007; Amos, Greaves, Nichter, & Bloch, 2012; Larsen, Otten, & Engels, 2009), the argument to smoke so that the person can lose or loses weight is presented by only one tweet in the sample:

“If you try to convince me to quit smoking after telling me how great I look now I've lost weight then you have already lost the argument..”

There are ten tweets categorised by this theme which also consist of content relating to the connection between losing weight and smoking, but these tweets are written as jokes or sarcastically about the idea of losing weight through smoking:

“@####### Cocaine and cigarettes: perfect combination to lose weight.:)”
8.3.2 Changing unhealthy co-behaviours

Within the co-behaviour tweets on smoking and healthy co-behaviour, the second most tweeted about theme with 153 tweets was the complete change in health co-behaviour where young people quit smoking and then started either eating healthier, doing physical exercise, or a combination of both. Research has shown that the perceptions of a healthy lifestyle and the knowledge of the health risks of smoking decreased tobacco intake in young adults (Gagné et al., 2015) and the sample members share these perceptions in Twitter content. The tweets in this category consist of changing health perceptions and the change in lifestyle the young people go through. Most of these behavioural changes start with a plan such as demonstrated here:

“Healthy eating, gym & no smoking starts today 😷”

“I need to stop saying and start doing! Motivated more than ever! #eathealthy #getfit #newjob #quitsmoking #driveagain #freshstart”

By far most tweets in this category are about the plans to do it and the success of their plans within the first week. Once the sample members tweet about being further along in the lifestyle change, there is a split between the types of tweets. First, there are those who tweet about feeling significantly better:

“Reaping the benefits in the gym since I’ve given up the fags. Body feels x100 better for it.”

“2 weeks at the gym :D 8 days not a single cigarette :D:D #happy”

Second, there are those that tweet about the difficulty and struggles that have arisen now that they have quit smoking and have started to diet and exercise:

“this whole going on a diet, being healthy, giving up smoking kind of thing. Yeah, really isn’t going to well”

“Dietering and cutting down on fags = angry emotions”

Some have argued that the reason people tweet about their success stories is that social media is seen as a place to show off how great life is (Zhao & Rosson, 2009). However, Twitter is also a place where people present more frustration than other types of social media (such as Facebook) (Harris et al., 2014; Jamison-Powell et al., 2012) and provides a setting to post about the struggle of quitting smoking. The sample members display both.
8.3.3 Unhelpful combinations

A number of tweets (36) in the smoking and healthy co-behaviour category were reflections from the young person on their health co-behaviour, and they seem honest about it being harmful to their health:

“Got to get out the habit of doing exercise, then going for a fag and then eating heavily buttered toast, defeats the purpose really”

“My worst habit is smoking a cigarette straight after the gym, defeats the object”

According to several studies, e.g. Amos & Bostock (2007) and Rodriguez & Audrain-McGovern (2003), physical fitness is a motive for not smoking as smoking affects the endurance of the body and physical exercise can be a good replacement for smoking in terms of needing stress relief. The tweets in this theme all have a reference to smoking affecting physical fitness and how smoking is unhelpful for proper physical activity:

“Was DYING earlier in gym... Defo stopping smoking jheeeezz”

“Haha! Need a fag after that workout! What's the point ai!”

Another common self-reflection originates from going to the grocery store. The young people either tweeted about going out for cigarettes and coming home with unhealthy food or contemplating how they spent all their money on cigarettes and unhealthy food:

“Go to the shop for tobacco and I come back with crisps, sweets and chocolate #suchachild”

“The money I used to spend on ciggs is now gonna be spent on fruit... So I'll be eating a lotta fruit! #ChangedMan”

Both examples illustrate how going to the store without buying the healthy option is associated with being immature. It seems that where personal health is concerned, healthy co-behaviours are considered responsible whereas smoking depicts the opposite and the content of the healthy co-behaviour tweets show that contrast.

8.4 Content of the multiple co-behaviour tweets: an in-depth focus

This last section of the in-depth focus on co-behaviours discusses the tweets with multiple health behaviours within the same tweet (e.g. smoking, alcohol and physical exercise). This number of tweets is small (105 or 7.8% of the co-behaviour tweets), and some have
already been touched upon in the previous sections. The following paragraphs outline the three themes that came up in the multiple health co-behaviour tweets.

**8.4.1 Combining unhealthy behaviours**

A common category consists of tweets containing multiple poor behavioural choices such as the combination of smoking with alcohol and unhealthy eating or no exercise. ‘Combining unhealthy behaviours’ is also related to emotion regulation as health risk behaviours all present a ‘quick fix’ against stress and general negative feelings. Likewise, the college students in the study by Nelson et al. (2008) who perceived more stress were more likely to smoke, be obese, and be physically inactive. These results, together with those of Little (2000) on the use of smoking and alcohol consumption for stress relief, indicate that it is common to refrain from healthy behaviours when the individual perceives tension.

The tweets in this theme, nonetheless, have mostly an upbeat tone and the pleasure of indulging in multiple unhealthy behaviours is seen in 27 tweets such as the following two:

“*Switzerland is amazing, weed, beer on tap, sh*t loads of food and everyone's been so nice. Class*”

“*Actually have the best friend ever ♥♥ when I'm down she brings me. Cider, chocolate and cigarettes #topgirl*”

However, not all of the tweets within this theme are positive, and the content of five tweets indicate the remorseful consequences of unhealthy behaviours such as the following:

“*My body hates me from too much drinking, smoking and eating #holidaycomedown*”

**8.4.2 ‘The whole health streak’**

Most predominant in this category were tweets in which the healthy co-behaviours were performed at the same time. Wiefferink et al. (2006) argued that when people want to quit smoking to become healthier, they are more likely to engage in healthy eating and exercising. The group consists of 42 tweets about performing the different health behaviours in combination. The sample members quit smoking, cut down on or quit alcohol intake, do physical exercise, and eat more healthily. Specifically, New Years’ resolutions, Stoptober and Lent were mentioned in eight tweets as times for attempting ‘the whole health streak’:
“New years resolutions - reach 7stone, quit smoking for good, study more, drink less fizzy drinks, save money...”

“Giving up cigarettes, caffeine, alcohol and crappy junk food has given me so much more energy. I love it!!! #stoptober #icanbreathe”

On the downside, changing lifestyle takes all the fun out of life according to five different sample members. Here are two of those tweets:

“How am I doing stoptober giving up drinking and eating healthy all at the same time. I'm in bits, I miss the naughty things in life”

“Miss food, miss normality, miss friends, miss drinking NOT missing smoking!!! #fedup #wanttofeelnormalagain #downday”

8.4.3 Comparing smoking health behaviours

Comparing health behaviours was a common theme among all the health co-behaviour tweets, and there were 86 tweets in total, i.e. 19 comparing multiple health behaviours, 51 comparing alcohol to smoking, 13 comparing healthy eating to smoking, and 3 comparing physical exercise to smoking. An important theme for the young people was the comparison between health risks behaviours and how that is (in their eyes) unfairly skewed towards regulating smoking. This dissatisfaction most likely relates to the extensive control on the sale of tobacco compared to the ease with which these young people acquire alcohol (Fuller et al., 2014; Parry et al., 2010; Tyler, 2015). The arguments that are given mainly focus on the lack of regulation of alcohol and unhealthy food compared to smoking:

“There should be severe health warnings on fast food and alcohol as there is on cigarettes, no?”

“how can Wales ban e-cigs and yet do absolutely nothing about it rising obesity and alcoholism problem? F*cking idiots.”

“The police should focus more on underage drinking than people smoking weed, i'm just sitting here chilling man”

Marijuana smoking regulation, especially, seems to be a sore point for young people. There are numerous adverse effects of marijuana smoking, but the marijuana-related tweets that are posted on Twitter diminish the perceived harm relating to this behaviour, and this positive perception is taken over by young people (Cavazos-Rehg et al., 2015, 2014). The awareness of what is healthy and unhealthy behavioural combinations are evident to the
young people, but they feel an unfair distribution of regulations that impede them to smoke.

8.5 Concluding remarks

The objective of this chapter was to examine the co-behaviours present in the smoking-related tweets through qualitative content analysis. This chapter is a continuation of the quantitative content analysis of Chapter 7 and provides in-depth scrutiny of the smoking-related Twitter content by how it is combined with other health behaviours, i.e. smoking, alcohol use, healthy eating, and physical exercise. The health co-behaviours in this sample of smoking-related tweets were examined by occurrence, gender and age, which in the previous two chapters have been shown to relate to smoking and Twitter activity. Here they were shown to relate to differences in how multiple smoking co-behaviours are expressed in the content of the tweets from the sample members. The occurrence revealed that a combination of alcohol and smoking is most by far most common. The gender component illustrated how women were more likely to add context to their tweets than men; a finding that is coherent with the findings of Lenhart et al. (2014) (among 800 12-to 17-year-olds in the USA where women were more inclined to share content on Twitter). Similarly, the age variation is consistent with the findings from Hughes et al. (2012); younger people are using Twitter more as an outlet for emotions than older young people. This showed that the content of the co-behavioural tweets is different with gender and age but that this might relate more to motives for using Twitter than smoking behaviour which is a reoccurring finding in these results chapters.

The sentiment score of the co-behaviours followed a corresponding pattern to that of Chapter 7 and illustrated that the individuals tweeted most positively about healthy co-behaviours in which they quit smoking. This result fits with success stories being most popular on Twitter (Zhao & Rosson, 2009). The negative sentiment scores are also an extension of the sentiment analysis in the previous chapter showing that the negative scores are highest when tweets are comparing health behaviours. It was seen that the young people felt an unfair focus on smoking regulation, specifically marijuana regulation which made the posts more negative.

The in-depth examination of the tweet content according to health behaviour has provided some general conclusions. The young people tweet mostly about smoking and a co-behaviour to resolve an unwanted emotional state (e.g. stress) and tweet about combining unhealthy behaviour for social purposes (e.g. consuming alcohol and cigarettes at a party
or buying them in for Christmas). Similarly, young people make the decision to become healthier (cut down on or quit alcohol, by exercising, or eating healthier) and gladly tweet about that too. However, most of these tweets are based on intentions and the first week in the changed lifestyle. After that, the tweets divide into positive tweets on how healthy they feel and negative tweets about how living healthily takes all the fun out of life.

Previous literature has shown that these health risk behaviours cluster in specific areas, for example, if people smoke, they are also more likely to consume alcohol, have an unhealthy diet, and exercise below the recommended guidelines (e.g. Mistry et al., 2009). Likewise, if people quit smoking and reduce alcohol consumption, they are also more likely to eat healthily and get enough physical exercise (e.g. Peters et al., 2009; Wiefferink et al., 2006). This chapter confirms these findings and adds knowledge on how these behaviours interplay with each other. From the content of the tweets, it appears that young people are aware of what is healthy and what is not, but the positive attributes of health risk co-behaviours frequently come out on top.

This chapter has increased our understanding of the social meaning of smoking by qualitatively examining the content of the smoking tweets that have a reference to other health behaviours. These findings show how the young people in the sample communicated their thoughts about smoking and emphasised how health co-behaviours are portrayed in helpful and less helpful combinations. To reduce the economic burden of unhealthy co-behaviours, the focus should move on from providing information about these unhealthy combinations to support the replacement of unhealthy behaviours with healthy co-behaviours, for example encouraging people to undertake physical exercise to relieve stress or enjoy a fresh orange juice on a sunny day instead of a cigarette and an alcoholic drink.

The next chapter is the last of the results chapters and centres on a subsample of 50 randomly selected sample members from the complete sample population to examine the wider online context of these smoking tweets with a qualitative analysis of the subsample’s entire Twitter feeds. This reveals the young people’s motives for using Twitter and the way these young people tweet about smoking in more detail.
Chapter 9. Smoking tweets in context

The purpose of this chapter is to provide an in-depth analysis of the smoking tweets embedded in the Twitter archives of a small selection of the sample members. As not every detail of a person’s life is shared on Twitter, the smoking content must have been noteworthy to communicate to their followers. This part of the results uncovers how these smoking-related tweets are situated amongst the content of other tweets to better understand why this smoking content was important to post. In the previous results chapters, the tweets were analysed in isolation and then connected to other smoking tweets by their content. By examining the smoking-related tweets in combination with tweets that are outside of the subject, the possible motives for, and the social context behind tweeting smoking messages can be explored further. Moreover, previous chapters have illustrated the general Twitter activity and sentiment scores of the smoking tweets, but there has not been an examination of the Twitter style (which combines the two) that possibly affects the smoking-related Twitter posts. Thus, the objective of this chapter is:

- to identify the wider context of smoking for young people as evidenced by their Twitter archive.

For this study, fifty sample members were randomly selected through a random selection routine available on SPSS (IBM Corporation, 2013) from the list of usernames used in this study (see Chapter 4 section 4.4 to read how the full list of usernames was created). The information of these individuals was placed in a separate table. In that table, the 50 subsample members were anonymised and given names in line with their gender (this table is shown in Appendix E). After that, all the tweets available from the fifty selected young people were collected and uploaded individually to Atlas.ti 7 (Scientific Software Development, 2013). Atlas.ti is a qualitative research programme that enables systematic coding of a large amount of data. This program provides a clear overview of the codes used in this analysis (e.g. presents all tweets with a marijuana reference and their content from all the 50 subsample members) and therefore, allows for a systematic analysis of discourse revealing meaning and patterns across different Twitter archives. The codes produced in Atlas.ti and the way they were layered with multiple codes to fit the objective of this chapter is shown in Appendix F.

This Twitter data is reviewed from two angles. First, the general Twitter activity of the 50 subsample members is examined to identify how these people commonly use Twitter so that a description can be made of how these individuals normally tweet and how that
differs when they tweet about smoking. This angle has a similar aim to the addition of the
general twitter activity in the time series of Chapter 7. Secondly, the smoking-related
tweets are examined within a framework of other tweets made in the surrounding 24 hours.
This approach is taken to find the context of when the tweet was posted and leads to a
better understanding of why these tweets were made.

The first part of the analysis consisted of examining the motives for using Twitter in the
Twitter archives (section 9.1). These motives were deduced by coding most strings of
tweets of the Twitter feeds of all the 50 subsample members into general codes such as
‘conversation’, ‘boredom’, and ‘sports commentary’ after which they were categorized into
motives (i.e. general interest and socialising) adapted from a study by Johnson & Yang
(2009). In the following section (9.2), the Twitter feeds of the subsample members were
further analysed through ‘AnalyzeWords’ (http://analyzewords.com/), an online linguistic
analysis program to uncover Twitter style. This program analyses Twitter feeds by
‘Emotional style’, ‘Social style’ and ‘Thinking style’ and was performed for each of the
subsample members. The variance in Twitter style is explored by smoking status and
gender.

After the Twitter usage analyses, the smoking tweets were identified in all the Twitter
histories (section 9.3). These markings of smoking tweets made it easier to discover their
direct tweet context, especially as all the smoking-related tweets, already had a code from
the ‘motived to use Twitter’. The smoking-related tweets and their context were analysed
systematically for each 50 subsample member. The analysis continued by examining the
smoking tweet’s fragments in groupings according to smoking status (i.e. non-smoker,
smokers, quitters and relapers), the setting (i.e. the discourse within the tweet) for
different smoking products and tweet content noted for its stigmatising discourse which is
discussed in section 9.3. To finish, the context (i.e. possible motives) of quitting smoking
and the Twitter interactions between this subsample population and The Filter Wales was
demonstrated in section 9.4.

A descriptive study of the fifty sample members revealed that in comparison to the total
sample population, they were reasonably representative (see Appendix E for the full table).
Through the API Twitter program, 113,569 tweets were collected across the subsample
containing 313 smoking-related tweets. Out of this subsample of 50 people, 16 individuals
were characterised as smokers, 13 were classified as quitters, 4 were categorised as
relapers, and the remaining 17 young people were considered non-smokers. This smoking
status is derived from a sequential read through of the smoking-related content per
individual (see Chapter 4 section 4.5.2 for more details) and throughout this chapter, the online smoking status is given with the examples from their Twitter feeds.

9.1 Motives for using Twitter

This section of motives for using Twitter aids in the understanding of the social meaning of smoking by examining, for each individual member of the subsample, the combination of why they use Twitter as this impacts the context in which the smoking-related posts are made. Incentives for tweeting were studied by Johnson & Yang (2009) in an online questionnaire of 242 people, and their results presented a number of social and informational motives people give for using Twitter. “Information motives included: get information (facts, links, news, knowledge, ideas); give or receive advice; learn interesting things; meet new people; and share information with others (facts, links, news, knowledge, ideas). Social motives included: have fun; be entertained; relax; see what others are up to; pass the time; express myself freely; keep in touch with friends or family; communicate more easily and communicate with many people at the same time” (p.17). Johnson & Yang provided thirteen possible motives for using Twitter. However, these motives do not seem to be combined completely logically as ‘relax’ and ‘be entertained’ read like information motives and ‘meet new people’ reads more like a social motive. Therefore, the distinction they make between what is classified as information motives and social motives is not adopted. Instead, their categories were adapted, and two subheadings (general interest and socialising) came out of it that better represent the tweets found in the sample.

9.1.1 General interest as the motive

Many of the motives mentioned above can be combined into ‘general interest’ (have fun, be entertained, relax, see what others are up to, pass the time, get information and advice, and learn interesting things) which is evident in the tweets. In other studies, Twitter is often used by people to look at what other people have posted and finding interesting tweets is an effective time-wasting activity (Livingstone, 2008; Morris, Teevan, & Panovich, 2010). This motive is found in the Twitter histories when individuals post a random string of tweets and retweets. Sarah showed this type of behaviour regularly in blocks of tweets such as these:

“Why can’t I just go to sleep. Takes me ages!!!!!”

“The thing I’ve been looking forward to all day has finally come! ......Aahhhhh much better”
“RT @#####: Dead ants give off a scent that tell the other ants they need to be carried away. The ant is then brought to a designated place”
“RT @#####: No matter how little money I may have, I will never write a Facebook status saying "someone lend me a tenner til Monday?” (Sarah, smoker)

These sequences of tweets in the subsample of 50 are challenging to deconstruct as they do not follow a logical structure and the tweets’ content can vary over a wide range of topics within a short time frame (between 10 minutes and 2 hours). According to literature, the motive of ‘general interest’ is a defining feature of social media and particularly Twitter, which is characterised by microblogs (D. J. Hughes et al., 2012; P. R. Johnson & Yang, 2009).

9.1.2 Socialising as the motive

The other motives given in Johnson and Yang’s research (2009) are ‘express myself freely’, ‘keep in touch with friends and family’, ‘communicate more easily and with more people at the same time’, and ‘meet new people’. These motives fall under the general motivation heading of ‘socialising’. Twitter users tend to follow others that have similar norms and interests as themselves and, therefore, socialising on Twitter is a form of reinforcing bonds with like-minded people in an online setting (M. S. Smith & Giraud-Carrier, 2010). Many of the posts from the sample members are conversations with others they know offline as well (Subrahmanyam et al., 2008).

Original tweets are posted by an individual to prompt or stimulate a response from their followers and start a conversation on Twitter (Hutto et al., 2013). Unfortunately, the Twitter gathering program outputs do not include the tweets sent in response to this sample of young people, and therefore half of the content of the conversation gets lost. Here is an example of Lucas who posted an update that interested one of his followers:

“Can tell I’m gonna have a massive spot on my forehead in the next couple of days and I can’t wait”
“@##### I shall take a nice photo for you when it finally appears”
“@##### I know I know” (Lucas, non-smoker)

Another example of tweets that would fall under ‘socialising’ on Twitter is for individuals to post live commentary on their Twitter account. These updates are not overly personal and relate mainly to an event that is taking place at the moment (such as a football match). At these times, young people share their opinions with their followers:
"City fans now "league means nothing" wasn't saying that last year
(…)
@#####1 for a premier league winning team buying, sinclair, rodwell and garcia says it all...
@#####1 true though isn't it...gotta make better signings than them 3 pair of pricks
@#####1 I'd say it's a good billion spent Ifa cup and a premier league :)
@#####1 @#######2 and still a better team than city” (Aaron, non-smoker)

The 50 subsample members apply the two core motives for using Twitter and reveal a bit of the reasoning for tweeting smoking-related content (i.e. to induce a response, socialise, and to share information with others).

9.2 Twitter style

After exploring the motives for using Twitter in the previous section, this section inspects what kind of style is present in the individuals’ collection of tweets through a linguistics program. Linguistic Inquiry and Word Count (LIWC), a program that uses science-based psychology in which certain personality traits are revealed through the number of words but most especially the type of words used in written language, is often applied in Twitter-based studies (e.g. Golbeck, Robles, Edmondson, & Turner, 2011; Hutto et al., 2013; Jamison-Powell et al., 2012). Contrary to the sentiment analysis on the individual tweets, this program analyses the entire Twitter archive of an individual and provides many more features of Twitter engagement. In this study, a free spin-off program ‘AnalyzeWords’ is used to uncover specifically Twitter Style. This program classifies the Twitter discourse for one individual by three style types: ‘Emotional style’, ‘Social style’ and ‘Thinking style’ and has been previously expanded on in Chapter 5 in Table 5.2. The derived style categories are analysed according to smoking status to uncover if smoking status relates to how the sample members tweet in general.

A high score in the Twitter Style LIWC means that the style category is present whereas very low says it is mostly absent. It was not possible to perform the linguistic analysis from the young people in the sample that have an inactive or ‘protected’ Twitter profile at the time of the Linguistic analysis (July 10, 2016). This resulted in a sample size of 40 (16 non-smokers, 9 smokers, 12 quitters, and 3 relapers) and an equal divide between men and women. The following sections explore the outcomes of the Twitter Style analysis. The tables that resulted from this analysis are found in Appendix F.
9.2.1 Emotional Style

The first category of Twitter style is emotional style and relates to the feeling a person puts in their tweets. This style is divided into four categories ‘Upbeat’, ‘Worried’, ‘Angry’, and ‘Depressed’.

The first of the emotional style categories is ‘Upbeat’ and refers to the number of positive words that are used in the tweets and positive interactions. The scores illustrate how most of the people in the sample are using positive words and sentence structures. The quitters are the most ‘Upbeat’ in their tweets and the two quitters that scored ‘very high’ are both women whereas the three people that scored ‘very low’ were men. Chanel and Jessica are the highest scorers in this upbeat category, and they post cheerful personal updates:

“Love these 2 so much ♥♥♥♥♥ (with a picture of her and two friends)” (Chanel, quitter)

“HAPPY BIRTHDAY TO THIS WONDERFUL WOMANY! THE WORLD BECAME A BETTER PLACE 18 YEARS AGO TODAY! MY PARTY PARTNER ♥♥♥♥♥♥♥” (Jessica, quitter)

In addition to the words, written sounds, punctuations and emoji’s are considered in the Twitter style analysis. As mentioned in Chapter 4 section 4.6.2, these additions make the intention of the post clearer for the reader, and in these cases, it illustrated an upbeat tone.

Anxious language and nervous questions are part of the score for the ‘Worried’ emotional style. The majority of the sample members fall between ‘low’ and ‘average’, but five quitters are most worried out of this subsample with a high score, and the very high-scorers are all women. Pippa is the highest scorer with 98% for this category, and this is one of her more extreme tweet strings that lead to this ‘very high’ score:

“we can't guarantee that we'll always healthy”
“my mom was healthy, almost far from disease, and now she's gone”
“home alone, thinking about mom, and now I feel terribly sad :'(" (Pippa, non-smoker)

Another woman portrayed the ‘Worried’ type of emotional Twitter style a little differently:

“can everyone just not”
“i feel like everyone is ignoring me”
“can't deal”
“literally going to have a breakdown” (Catherine, non-smoker)
The example of Catherine illustrates how the lack of interaction on Twitter made her more nervous. Linking the ‘Worried’ Twitter style to the motive of socialising, the high scorers in this section are looking for engagement from followers.

The third emotional category defines the ‘Angry’ measure which is produced by the use of caps for full tweets, the use of hostile words, and writing a lot about ‘you’. Twitter is a social network site that is used to vent frustration more often than other ones (D. J. Hughes et al., 2012; Kwak et al., 2010) which is exemplified by the score where 55% of the smokers scored ‘very high’ on the ‘Angry’ category in the emotional style linguistic analysis which is far more than the other people in the sample. Interestingly, none of the young people in the sample score ‘very low’. In the subsample, often the ‘angry’ tweets are about everyday life annoyances. These are tweets that mention annoyance or irritation:

“Unpopular opinion but I HATE EASTER EGGS I CAN BARELY MANAGE CHOCOLATE AT ALL” (Diana, smoker)

Nothing worse than the bath water going cold ffffss” (Zachary, smoker)

For the ‘Depressed’ emotional style category, tweets that are sad, full of melancholy and inward-looking are measured. Moreover, the score is created through the number of self-references and the frequent use of depressive words. In the ‘Depressed’ analysis, no one scored ‘very low’ which is similar to the ‘Angry’ measure. Interestingly, one-third of the smokers scored ‘very high’ which is another emotional style measure that is typically apparent in smokers. Two examples from the Twitter feeds of the subsample members that fit with a high ‘depressed’ score are presented here:

“Second night in a row ive got to bed crying yay. F*ck men” (Wendy, non-smoker)

“some times I feel I have mental issues but then I'm just like stfu n cheer up u loser n i feel 10xs worse”
“for the first time in my whole life I just looked at myself in the mirror and thought my skin looked good”
(....)
“I'm so bored :((((“
“I'm just gunna lie in bed n watch doctor who till someone decides they want to go out n play” (Jessica, smoker)
These two examples demonstrate depression but not in a way that they would want help or at least there is no indication of it in the following tweets. Rosenquist, Fowler & Christakis (2011) studied depression on social network sites, and they argued that these posts are reinforcing other posts about depression in homogenous networks. The originators are looking for compassion from others and not professional help (Rosenquist, Fowler, & Christakis, 2011). Similar to the other emotional styles, it is a call for interaction with their followers and, as Yang & Johnson (2009) would argue, fits with the use of Twitter to ‘express oneself freely’.

9.2.2 Social Style

The second Twitter Style category from the AnalyzeWords linguistics program is a measurement of the social markers individuals put in their tweet content. The Social Style measure refers to how they engage with Twitter and especially how they interact with other people on Twitter. The categories consist of ‘Plugged-in’, ‘Personable’, ‘Arrogant/Distant’, and ‘Spacy/Valley Girl’.

The ‘Plugged-in’ measure relates to the connection Twitter users make with others. It scores social words such as ‘party’ and particularly measures the number of references to other Twitter users (i.e. @########). The analysis demonstrates how the majority of the sample population scores average on this measure, and there is hardly any distinction between the people in the subsample. Nicole is a high scorer, and her Twitter feeds are filled with updates like these:

“Tonight’s agenda...House, Fringe, Revision and E-skulls...#partyonwayne”
(Nicole, quitter)

The highest scoring man in the ‘Plugged-in’ category is Quentin, and he often uses Twitter to congratulate people:

“HB @##### Cant wait for 2moro night, looking forward to your dirty pint!!
#cokefantaandredbull. #comeonitsyourbirthday.” (Quentin, non-smoker)

Being ‘Personable’ on Twitter indicates that the Twitter user cares about other peoples’ wellbeing and posts about their own uncertainties through, for example, asking questions and often referring to others. All the selected individuals scored high on the ‘Personable’ measure, but the smokers and relapsers stand out as being most personable. A high personal score is achieved when people use Twitter as an online diary. Most of their tweets refer to their emotional state and opinions at the time, and Stephen provides an example:
“I rather dislike being dragged out to buy school uniform #boring”
“I regret ever talking to someone.. wow”
“Well tonight has been utter crap. Now time to lay in bed and try and sleep”
“I have had it with women for now” (Stephen, non-smoker)

But not all the high scorers in this category post about themselves. One of the top scoring women is Chanel who has conversations about well-being with others:

“@###### what u done babes”
“@###### bit better. still bit wingy.. what u doing tomorrow?”
“@###### well mollie in nursery. and im off work with jack. so guna be bored..
jack cnt go nursery.”
“@###### ok thanks.. very much appreciated.” (Chanel, quitter)

The ‘Arrogant/Distant’ measure rates an impassive way of tweeting and a lack of self-referencing in tweets. The ‘Arrogant/Distant’ category had a low score overall, but half of the people scoring average are quitters and represents all but two of the ten quitters. The highest scorer is Gary with a total score of 76% on the ‘Arrogant/Distant’ measure. He tweets mostly live commentary of football matches but alongside that he tweets about the failings of others:

“It's genuinely an absolute RULE that the people who send abuse can't spell
"you're". It's like a gang sign” (Gary, non-smoker)

The ‘Arrogant/Distant’ social style is not popular with the subsample members which is not remarkable as people want to bond with their followers not distance themselves from them (M. S. Smith & Giraud-Carrier, 2010).

The ‘Spacy/Valley Girl’ measure relates to the necessity of recounting the newest stories. Moreover, the score is based on the use of abbreviations such as LOL (laughing out loud) and ROFL (Rolling on the floor laughing) as well as an overuse of punctuations. Over half of the smokers and 43.8% of the non-smokers had a ‘high’ score on the ‘Spacy/Valley Girl’ category. Most of the people score ‘average’, but women scored higher than men. Abbreviations are a preferred method of writing for the young people and are taken up in the ‘Spacy/Valley Girl’ category. They are used in tweets to speed up the writing time, and the examples given below show that they are used in interactions on Twitter:

“@###### probs totes done this many a time” (Nina, smoker)

Translation is @###### Probably, I have totally done this many a time.
The Social Twitter style reveals that the sample members are social on Twitter and tend to ‘overshare’ if the tweets were intended for public view. The young people seem to forget that anything posted on Twitter is publicly accessible. Mao et al. (2011) argued that the privacy leaks are high and most people do not realise what danger they put themselves in by using Twitter is such a personal way.

9.2.3 Thinking Style

This section of the linguistic analysis refers to the level of thought that went into the writing of the tweet or the decision of retweeting someone else’s tweet. The categories in the Thinking Style are ‘Analytic’, ‘Sensory’, and ‘In-the-moment’.

The ‘Analytic’ score of the subsample members describes the way people create sentences. In this linguistic analysis, the non-smokers and smokers get highest scores in the ‘Analytic’ category. A high score indicates that people make full sentences with proper punctuations and they use full sentences like these examples:

“@##### you should, the Exeter mafia has kidnapped me and are requesting a £500 ransom delivered to my house if you'd be so kind” (Stephen, non-smoker)

“Can we appreciate that sometimes when I don't reply it's because I'm having some serious life issues and just want to think” (Diana, smoker)

Especially when Twitter is used to socialise and interact with others, the speed of posting a tweet is important which makes grammar less imperative. So, depending on their motive for tweeting the analytic score fluctuates within the Twitter feeds.

The ‘Sensory’ measure relates to the censoring of the originator and how much careless opinions are given. For example, the lack of swear words in the tweets gets individuals to become top scorers. The ‘Sensory’ scores divided by smoking status reports that all three relapsers in this analysis score ‘very high’. Only the non-smoking group has a more diverse score and the two individuals that scored ‘very low’ are both quitters. Six members of the sample score very high (>95%) in this category. The tweets from their Twitter feeds consist mainly of non-offensive updates and opinions. Moreover, these six people score quite low on the ‘Depressed’ and ‘Personable’ categories and here are two:

“Can’t get enough of Peaky Blinders! #classseries” (Olaf, quitter)

“Missing my woman millions :( after spending a week with you is still not enough. #sadegg xxx” (Uma, smoker)
The ‘In-the-moment’ measure scores how ‘up-to-date’ the Twitter user is. Being ‘up-to-date’ manifests itself by the use of hashtags and the quick retweet or reply to original posts. There are no apparent differences in the ‘In-the-moment’ category except that the four low scorers are quitters and non-smokers. Hashtags are used by all the sample members, but the women tend to use them a lot more. These hashtags are used to inform their followers in a transparent way about their opinions and feelings (Tsur & Rappoport, 2012) as seen in the tweets above by Olaf and Uma.

Another ‘In-the-moment’ indicator are retweets. Most of the retweets are tweets from the news, so-called ‘factbanks’ (unverified facts on Twitter profiles), celebrities, or organisations the young people are following. These tweets pop up on their Twitter wall, and while Twitter users are scanning through over all the tweets, some are interesting, and they decide their followers should see that too (Macskassy & Michelson, 2011). Pippa and Taye, two high scorers, present two examples:

“RT @#######: #Taurus is very stubborn.” (Pippa, non-smoker)

“RT @#######: Benefits shakeup aims to force more disabled people into jobs”
(Taye, smoker)

The Thinking Twitter Style relates more to the forms of the tweets and the individuals in the subsample score above average in this area. This indicates that they think about their Twitter behaviour and online persona.

In conclusion, the subsample members score high on the ‘positive’ categories. Smokers scored highest in the ‘Emotional’ style which indicates that this group uses Twitter more as an outlet for emotion than people with another online smoking status. The smokers also scored highest in the Social Twitter Style indicating that those styles are often closely linked to each other; if the person exposes a lot of emotions on Twitter, they use this platform for more social interactions too. The Thinking Style was more evenly divided by the members with different smoking status and besides did not seem to be connected to the other categories. All the Twitter Styles together indicate that the young people in this subsample use Twitter to extend their social life into an online platform and to bond with people with similar norms and interests.

9.3 The smoking-related tweets and context

Following on from the Twitter style of the subsample members, this section concentrates on the presentation of the smoking-related tweets and the posts linked to the smoking
activity. Out of the tweets that were collected from the subsample (113,569 tweets via the API program), 313 (0.3%) were related to smoking. All these tweets were analysed systematically and the smoking-related tweets and the other tweets in a 24 hour period were identified within the Twitter archives on Atlas.ti.

The context of the smoking tweets was scattered between random and orderly strings of tweets. Meaningful tweets that provide information about the everyday lives of the young people were often interspersed amongst tweets that were not so useful. These less useful tweets tended to contain texts that did not provide information on the context within which the smoking tweets were made. If the smoking-related tweets were not the only updates of that day, they were regularly accompanied by a random sequence of tweets which can best be linked to the motive of using Twitter out of ‘general interest’ from section 9.1.1 earlier. Paige demonstrates an example of rambling with these four tweets posted consecutively within two hours that presented no specific indication of the context of smoking:

> “Hoping tomoro at work will be a good night!”
> “One act of random kindness at a time”
> “Can’t believe it’s been a little over 3 months since I’ve quit smoking #neverthoughtidseetheday”
> “I really need a notepad and good pencils #wanttodrawagain” (Paige, quitter)

Less often tweets around the smoking posts were connected, an example being a negative emotion that is building up. As was mentioned in Chapter 8 and the literature (e.g. Little, 2002), smoking is used for stress relief. As illustrated in the ‘Emotional style’ (section 9.2.1), the smokers in the sample are a lot more vocal about emotions on Twitter, and following string of tweets shows how smoking appears as an emotion regulator:

> “Just googled the marking scheme for the past paper we were set up”
> “I am going to fail biology 100% don’t give a shit”
> (….)
> “Only 100 cigarettes can fix this” (Diana, smoker)

> “OH MY F*CKING GOD. #Pornhand”
> “I might have to sell all my clothes on depop to live this month”
> “everything's stressing me out”
> “need a big joint n a bottle of wine” (Jessica, smoker)
These two tweet blocks demonstrate the potential richness of using the full array of Twitter posts and the associated potential in social research. Single tweets about smoking fail to reveal social context, but in-depth mining of any one person’s Twitter traffic can help reveal elements of the setting within which the smoking activity takes place. This observation also highlights the methodological challenges associated with data mining and getting the full potential out of such big datasets.

9.3.1 Smoking patterns on Twitter

The smoking patterns of the subsample members become visible when the entire Twitter history is looked at in chronological order. This was previously done to create the online smoking status variable, but this time the smoking-related tweets are examined for commonalities within the online smoking status groups (i.e. non-smokers, smokers, quitters, and relapsers).

The non-smokers (17 individuals) are categorised as such as they posted nothing in the sense of having a smoking habit. Their smoking-related tweets concerned others smoking:

“Chinese toddler seen smoking in street http://##########” (Gary, non-smoker)

“hey man, do not smoking beside a woman please!” (Pippa, non-smoker)

For the non-smokers, the maximum number of smoking-related tweets within their Twitter feeds was three, and as they are not smokers, there is no smoking pattern to be found.

The smokers’ group consists of sixteen individuals that referenced a smoking habit in their Twitter histories. The label ‘smoker’ is given to any of these young people in the sample if they tweeted about smoking themselves. Three of the smokers successfully quit smoking tobacco but are still smoking marijuana or enjoying shisha which still constitutes being a smoker. Seven individuals with the ‘smoker’ status smoked marijuana in addition to their tobacco such as Taye and Zachary show:

“@###### fags is the big one for me mate, I'm a smoker of both, but I don't want to give up weed.” (Taye, smoker)

“I need to quit smoking man”

(...) 

“@###### hahaha my cigarettes have got like twelve million ounces of WEED IN EM BRUV GET ME FAM”

(...)

189
“When that weeds so WHACK I pull out DA fanny PACK http://######”
(…)
“Quitting smoking was by far one of the best decisions of my life!!” (Zachary, smoker)

The literature on youth smoking revealed that two-thirds of Welsh youth smokers desire to quit tobacco (ASH Wales, 2011). This was seen in the Twitter feeds of the smokers too as, even though they are tweeting about smoking, nine out of the sixteen smokers in this sample posted a tweet about the prospect of quitting smoking including Usain:

“My dad is constantly telling me to quit smoking then goes and buys me a pouch of gold leaf like son make ur mind up” (Usain, smoker)

The ‘I should quit’-tweets from the smokers showed dissatisfaction with their current smoking habit but not enough to change. Similar to the findings in Chapter 8, to become healthier by quitting smoking, these goals did commonly not go further than intentions. It seems they are placed online purely to instigate a response from their followers.

Thirteen out of the fifty young people in the subsample were classified as ‘quitter’. A notable find, when specifically focusing on the discourse of quitting, is that all the quitting attempts from the subsample related to quitting tobacco. The smoking tweets from the quitters were more plentiful than in other smoking status group, and this is likely due to quitting smoking being an achievement that appears well on social media. Moreover, it can be tweeted about multiple times as Paige exemplifies:

“5 days no cigarettes:D:D”
(…)
“Three weeks and a day without a cigarette #ontherightpath #nocravings”
(…)
“Two weeks from tomoro it'll have been two whole months cigarette free! #icanbreathe #icantastethingsagain”
(…..)
“Two whole months today cigarette free:D:D”
(…)
“Can't believe it's been a little over 3 months since I've quit smoking #neverthoughtidseetheday” (Paige, quitter)
Not all quitters posted multiple tweets about their successful quitting attempt. It occurred three times that the smoking status as ‘quitter’ was based on a single tweet stating the individual has quit smoking some time before:

“Has given up smoking....God be with you all #latestoctober” (Nicole, quitter)

“Can't believe I've quit smoking for like 3 months this month haven't even bothered me if I'm honest” (Yara, quitter)

The quitters in the subsample post tweets in an upbeat manner as could be seen in the Twitter Style. These ‘having quit’-tweets are filled with an uplifting tone and self-acknowledgement of achievement.

There are four individuals in the subsample categorised as ‘relapsers’. To tweet about failing is a little more common to express on Twitter, but social media is largely reserved for success stories and ‘funny’ updates that ‘need’ sharing (Hutto et al., 2013). This section is divided into two group; two individuals who first tweet about having quit and later tweet about smoking again (Emilia and Yvonne), and two individuals who explicitly mention their failed attempts (Zander and Kevin). For the first group, the discourse is not different from the smokers or quitters as they are positive about their quitting attempt and later tweet about smoking:

“foul mood feel like having a fag ffs #beenaweek #quitting”
(....)
“@##### I'm going for a fag x” (Emilia, relapser)

The other two relapsers are explicit about failing which suggest a ‘pity’ post to receive attention from their followers and ‘express themselves freely’ for socialising purposes (see 9.1.2 earlier):

“might aswell start attempt 999999999999999.1 of quitting smoking”
(....)
“Can't wait for my vape to get delivered, I'm hanging asf when I need a fag ?? time for quit attempt 1001 is it?” (Kevin, relapser)

This group of relapsers showed a gender divide which could be related to women being less likely to share negative content on Twitter, but a sample of four is too small to make any conclusions. It is quite possible that a few quitters in the sample relapsed as well but did not mention smoking later on in their Twitter feed.
9.3.2 The context of different smoking products

The context of the smoking tweets is highly dependent on the type of smoking product that is used. As illustrated in Chapter 7, there are different activities, times, and sentiment related to the various smoking products and the analysis of discourse on the subsample reports similar results as the following paragraphs will illustrate.

Whenever the fifty young people tweeted about tobacco smoking, they mentioned it ‘in passing’ as tobacco does not seem to be special enough to gain the required effects of online engagement. A central driving point for the unimportance was that tobacco is consumed in isolation more often than the other products (Myslín et al., 2013). Interestingly, tobacco was tweeted about merely if it had a specific place in the story of the tweet. The following examples indicate that tobacco is mentioned while another (more noteworthy) event is going on:

“I really need cigarettes but too afraid to walk around this town to get some #cold#crying#stranded” (Paige, quitter)

“Getting pissed off far too easily today. Time for a ciggy and tea break. Deep breaths....” (Benedict, quitter)

Eleven sample members mention marijuana in their Twitter feeds, and they mostly tweeted about smoking it. Thompson, Rivara & Whitehill (2015) found similar results in their content analysis of marijuana Twitter chatter where 54.9% of their collected tweets related to personal use. The tweets in this section show that marijuana has become more ordinary and available in their lives, and the young people in this subsample do not seem to have any problem with accessing it:

“wonder when @##### is going to have a sick weed infested house party...Saturday would be good” (Victoria, smoker)

“It’s so weird how smoking weed is such a normal thing these days, it’s more surprising to find out if somebody doesn’t smoke it :/” (Lucas, non-smoker)

Specifically, in the content of marijuana referenced tweets, many tweets are retweets or ‘copied’ from other people such as the following example:

“the world turning, the weed burning.” (Zander, relapser) lyrics from the song ‘the race’ by Wiz Kalifa
Thompson, Rivara & Whitehill (2015) found 15% of the marijuana-related tweets were pop culture references in their content analysis of 36,969 marijuana-related tweets from people in the US. They did not, however, attempt to delve into this observation and explain why these Twitter users tend to link pop culture references to marijuana smoking. A possibility is that the young people want to post something ‘cool’ but could not come up with the words for it themselves or it could show an association with the reggae culture or a specific pop group. Besides the pop culture references, several retweets of marijuana-promoting Twitter profiles were present in the texts of the subsample members such as the following:

“RT @#####. The growth of cancer cells can be slowed down by the consumption marijuana.” (Yara, quitter)

Tweets about marijuana ‘facts’ are popular to retweet for young people to showcase their stand in the debate about marijuana legalisation (Cavazos-Rehg et al., 2014). So, even though most tweets are about smoking marijuana products, a large portion of the tweets concerns other topics more closely related to pop-culture and online presentation.

As mentioned in section 7.1 in Chapter 7, the e-cigarette use in the sample refers to quitting tobacco smoking and not in its own right as was concluded in other literature, e.g. Goldstone et al. (2016) and Moore et al. (2015). In the analysis here an additional observation is made and refers to the difference between smoking tobacco and e-cigarettes. Several tweets refer to the complications with smoking e-cigarettes instead of tobacco:

“Literally gonna go all day tomooro without my ecigs #deadbatteries #shittysundaymornin #storyofmylife #gottagetacharger” (Paige, quitter)

“Running out of juice and being closer to a corner shop with cigarettes than a shop with e liquid is the worst #willpower” (Kevin, relapser)

Shisha smoking is occasionally referenced in the tweets and is mentioned by people who are not essentially in favour of tobacco smoking. The few individuals that mention shisha smoking do not perceive it as harmful or even identify it as a type of smoking product. Freddy and Uma, for example, mentioned:

“Smoking just ain’t me and never will be #blacklips”

(.....)

“If I had a hooka to end the night it would be perfect #happy4th” (Freddy, smoker)
“Watching Miranda on the way home on the train. Group of steamers smoking a spliff. I feel sick omg too much smoke :(["

(....)

“Wish I brought my shisha back from Bath” (Uma, smoker)

The overall impression from these tweets is that shisha does not belong in the same category as tobacco, marijuana, and e-cigarettes. In a systematic review of motives and beliefs about waterpipe smoking by Akl et al. (2013), it became evident that there is a common misconception among young people that shisha smoking is less addictive and less harmful than (normal) tobacco smoking as there is a lack of anti-shisha regulation and lack of knowledge about the health risks.

9.3.3 Stigmatising content

This section on stigmatising content was not anticipated when the analysis of discourse was planned and what follows has arisen from a largely inductive approach. The notion of stigmatising content appeared as the coding process proceeded and it was soon apparent that the concept of ‘stigma’ was such an essential element for the understanding of the social meaning of smoking that it was added to this chapter to reveal a deeper understanding of how tobacco control impacts young people. Stigma is seen in various ways in the smoking-related Twitter feeds of the subsample population. As was elaborated on in Chapter 3 section 3.2.1, smokers feel stigmatised especially after the smoking bans came into place (Parry et al., 2010; Ritchie et al., 2010b). In the Twitter feeds of the subsample members, mainly the non-smokers post negatively about the smoking habits of others:

“Sat next to a car in maccies and it has 2 parents and a child in it and the parents are smoking weed what the f*ck is wrong” (Rebecca, non-smoker)

“If pregnant women get cash incentives to stop smoking .. it will be a joke!! You think doing it for your unborn child would be enough! #nhs” (Francis, non-smoker)

Interestingly, this stigmatisation is also visible in smokers who marginalise others and how smoking has become an unpleasant activity to endure:

“Nothing worse than seeing someone pushing a pram and smoking #uch”
(Isabella, quitter).

“nothing more gross than smoking. It smells, looks and tastes rank! Hate people blowing it in my face! #wheresthecleanair” (Nina, smoker)
Both cases exemplify the originators of the tweet seeing themselves as ‘considerate smokers’ (described by Ritchie, Amos & Martin, 2010b) as those smokers have consideration for other people and are, therefore, ‘allowed’ to judge the smoking behaviour of others who are not. However, it does not stop at marginalising others as Victoria and Benedict expressed self-stigmatisation in their Twitter histories:

“Cannot, will not, haven't stopped spewing”

“I was smoking a fag, a bloody fag! Out the minibus window. I've changed, i barely recognise my self #shameful” (Victoria, smoker)

“Standing in the car park at work smoking an electronic ciggy just seems wrong @##### I can but feel bad! Silly really :P Still, 26 days so far :)(Benedict, quitter)

According to a systematic review of tobacco use and self-stigmatisation by Evans-Polce et al. (2015), tobacco control policies are presenting smokers with guilt and stress. The policies only affected smoking behaviour to a small extent, but most of the time it made the smokers feel bad about their smoking behaviour which can lead to severe mental health problems.

9.4 Quitting attempts and contact with The Filter Wales

This PhD study was undertaken in collaboration with the smoking cessation organisation The Filter Wales (which has the aim to help young people quit smoking). One of their methods is to interact with young people across Wales through Twitter (see Chapter 3 section 3.4 for more details on The Filter Wales program). This section explores the Twitter narratives surrounding quitting attempts that individuals provide on Twitter and how they interact with The Filter Wales partly to understand the social meaning of smoking better and partly to illustrate the context of The Filter Wales’ Twitter engagement. These segments are combined because most of the interaction of the Filter with the young people concerns quitting smoking but the Filter Wales social media team does not necessarily know the context of those ‘quitting smoking’-tweets. Therefore, this section begins with examining the tweets prior to the quitting attempt to uncover the motivation to quit smoking.
9.4.1 The quitting attempts narratives

The motives for quitting smoking could be uncovered by analysing the tweets preceding the ‘quitting smoking’-tweet of the individual in the sample. A quitting attempt often indicates a ‘rite of passage’ where something in their life changes (e.g. becoming a parent) and smoking no longer fits in with that lifestyle (Laurier et al., 2000). Here are two examples of how the posts about quitting smoking were preceded by tweets that offered an indication of motivation:

“Zumba timeeee 💃👙💪”
“Zumba done fitness done but I refuse to go on a jog in this weather”
“and I need new runnin trainers anyway”
“managed to do 59 crunches without stopping. feel so much more healthier/fitter since all this working out!!! #zumba #diet #exercise”
“Now I need to quit smoking...” (Emilia, relapser)

“Run later in prep for Mondays fitness test!!! #workinghard “
“To be fair I'm actually loving just chilling out tonight! Can't be assed to be drinking! #wastedmoney”
“@###### to right there don't think I'm gonna be going out until maila now anyway only 2 weeks away #bestgetsaving”
(....)
“Day 2 of no smoking.... Ohhhhh sh*t! Gonna be a long as day!” (Olaf, quitter)

However, most of the quitting smoking tweets did not have proceeding tweets to explain the incentive for the quitting attempt. Moreover, as mentioned in 9.3.1 previously, some ‘quitters’ only posted one smoking-related tweet with the information that they quit smoking. The individuals would most likely have had motives to quit smoking, but they did not post these on Twitter.

9.4.2 Contact with The Filter Wales

The Filter Wales is a youth-dedicated smoking-cessation organisation that contacts young people through Twitter as part of their campaign. When the Filter social media team finds tweets from young people about quitting or thinking of quitting smoking, they reply with words of encouragement and often a link to the Filter website. The website contains information on how to quit smoking which will hopefully make the quitting attempt more successful. Here are two examples of that type of interaction:
Isaac (quitter): “Haven't had a cigarette all week ether!”
The Filter: “Well done @#Isaac# you can do it stay strong, we are here if you need support http:#######”

Yvonne (relapser): “Day 1 no smoking.. HOLLAAAA”
The Filter: “How's today @#Yvonne##? First days are hardest, conquer those; then it's mind over matter https:#######”

Isaac and Yvonne did not respond to the tweet from the Filter, and this was very common with young people according to the Filter social media team. Four people from this subsample of fifty did respond to the encouragement of the Filter and here are two illustrations of those interactions:

Nicole (quitter): “Has given up smoking....God be with you all #latestoptober”
The Filter: “@#Nicole# it's never to late to give up smoking! Good luck :)”
Nicole: “@thefilterwales thanks :) I've got one of those e-cigarettes which is fantastic for me :)”
The Filter: “@#Nicole# glad you've got something that's helping you thru! Cold turkey can be tough. Good luck :)”

Maxwell (quitter): “@###### I have been cigarette free for 3 months now. And you said it wouldn't last!! #InYourFace”
The Filter: “#lovethistweet @#Maxwell# how have you stayed quit?”
Maxwell: “@thefilterwales just will power. Electric cig for first week then chewing gum and exercise. #FeelingTheDifferenceNow
The Filter: “Sorry for late reply @#Maxwell#. That's great, will power is a huge factor and motivation #feelthepower”

The Filter always makes sure that they respond to anything the person is tweeting with a personalised reply. This was considered a key activity which increased the reach and success of the campaign according to Meek, Hurt & Grant (2015) in their evaluation of the entire campaign. Moreover, other literature has shown how interaction and encouragement are most effective for individuals who want to quit smoking (e.g. Brown et al., 2014; Laranjo et al., 2014; Paay, Kjeldskov, Skov, Lichon, & Rasmussen, 2015; Ramo, Hall, & Prochaska, 2010; Richardson, Green, Xiao, Sokol, & Vallone, 2010). Besides encouraging smokers to quit, the social media team retweets posts to emphasise that young people are being heard about their perceptions of smoking:
“RT @#Quint#: cant believe smoking in pubs and clubs was once legal”

“RT @#Tess#: Group of men smoking weed outside Toys R Us. Standard.”

The interactions between The Filter Wales and the young people are not representative of who will be successful at quitting smoking instead, these interactions explain more about the sample’s Twitter Style and how willing they are to interact online with people they do not follow on Twitter.

9.5 Concluding remarks

This chapter aimed to provide context to the smoking-related tweets of a subsample of fifty young people in this study through analysis of discourse. Qualitative analysis is a time-consuming approach, and, therefore, it was important to take a smaller group to provide an in-depth understanding of the social meaning of smoking that can be found in tweets. This randomly selected group of fifty provided valuable information that enabled a fuller understanding of why the sample was tweeting certain smoking-related content and particularly, why they were keen to give up or not to give up smoking. The tweets surrounding the smoking-related tweet provided more information about the individuals’ life, concerns and motives than the smoking tweet alone.

It was important to uncover the motives and style of Twitter by illustrating the types of tweets that are placed on this online platform. The two overarching motives for Twitter activity are ‘general interest’ and ‘socialising’ and provide insight into the purposes of using Twitter and how smoking tweets fit within this setting. General interest resulted in random strings of tweets that could not provide real context to the smoking-related tweet. However, the motive of socialising provided more useful context and showed how smoking-related tweets were posted to relate and engage with others. This social motive is an important finding enabling a better understanding of how the smoking tweets (and other tweets) were made to provoke a reaction, and their content was designed for that purpose.

The way the subsample members use Twitter was also analysed by a program called ‘AnalyzeWords’ that illustrates Twitter Style and presents an indicator of how individuals use Twitter to create their online persona. The Twitter style of these young people suggested that original tweets were made during emotional situations, to establish social relations, and to a lesser extent report analytical thinking. As these Twitter styles were divided by smoking status, the different online personas could be analysed which revealed that smokers have a general Twitter style that is more emotionally negative and quitters
were the most upbeat and social. Murnane & Counts (2014) found similar results in their study on the Twitter archives of quitters and relapers, where, overall, the relapers used less positive language and were less sociable in their tweets than quitters. Although a causal association could not be made, it might help organisations, such as The Filter Wales, to better understand their target audience and produce the best strategy to interact with them, e.g. smokers are commonly more emotional Twitter users and the Filter could tap into that Twitter style (even more).

The patterns of smoking-related tweets made by the individuals illustrated according to smoking status showed how most smokers actually wanted to quit smoking but stopped at tweeting about the intention. This content induces a response from their followers and might be linked to the results from Uppal et al. (2013) where smokers ‘ought’ to want to quit but lack proper motivation. Similarly, for quitters, there were not that many ‘quitting smoking’ tweets that contained motivation but just excitement of the achievement to induce a response. Likewise, the interaction with The Filter Wales differed from person to person. The context of the tweets most likely relates to Twitter Style and motives to use Twitter in combination with the likelihood that young people make a quitting attempt unassisted as seen in the studies of Bancej et al. (2007) and Berg et al. (2010).

The different smoking products serve different purposes in the lives of young people; tobacco and e-cigarettes smoking seem uninteresting and more personal (e.g. mentioned as an achievement of having quit smoking) whereas marijuana and shisha smoking tweets were outward facing and smoking was mentioned in order to socialise and bond with others through Twitter (e.g. retweeting positive marijuana messages to show support for more lenient regulations). These findings are similar to the results from the time series analysis in Chapter 7 showing that tobacco and e-cigarettes content is more shared during the day whereas marijuana and shisha content is more often shared at ‘social’ times.

An important theme that came up in the analysis of discourse was that subsample members were negative about their own smoking habit and in a way marginalised others and themselves by posting marginalising content online. The high awareness of the social disapproval allowed non-smokers and smokers to critique others smoking and increased perception of having to self-criticise their smoking behaviour.

This chapter has highlighted that while young people share a lot of personal information on Twitter, the smoking behaviour context is incomplete. Smoking tweets have connections to other tweets of the same period, and in some cases, a pattern can be found. The young
people were inclined to use Twitter extensively and tweet about smoking-related topics, but these tweets were mostly in-line with their Twitter Style and posted to induce a response from their followers.

The next and final chapter outlines the conclusions of this thesis. This chapter will indicate the broader importance and contributions to knowledge that is offered by this study and expand on the possible themes and areas for future research.
Chapter 10. Conclusion

This thesis contributes to the body of existing work on understanding youth smoking persistence and the complex interactions between anti-tobacco movements and the smoking perceptions held by young people. As seen throughout this study, most of the existing research apply traditional methods of data collection, usually in the shape of surveys, interviews or focus groups aiding in our understanding of the social meaning of smoking. These traditional studies rely either on a large body of data gathered via observational approaches with relatively little detail on patterns or more detailed data derived from a relatively small number of cases to uncover young people’s perspectives.

While some studies have focussed on smoking references obtained via Twitter posts, Twitter data and sociological insights have not hitherto been combined to the extent attempted here. Thus far, few researchers have used social networking sites as research material, but this data can provide a significant amount of qualitative data and deliver uncensored perspectives on smoking. Moreover, Twitter data contains additional information such as geolocation and date of the smoking-related content which, as demonstrated in this thesis, proved useful in examining where young people (interested in smoking) originate and when smoking is important for them. Crucially, this thesis demonstrated how, by using Twitter as a research tool, a new layer could be added to understandings of the reach of a smoking cessation organisation and the social meaning of smoking portrayed by young people on Twitter. The overall aims of this study were;

1. to understand the reach of the Twitter element of a social media campaign (The Twitter element of The Filter Wales).
2. to assess the text content of tweets about smoking in order to understand more about the social meaning of smoking and health risk co-behaviours.

This chapter outlines how these aims have been met by summarising the results of the thesis (section 10.1) and underlining the contributions to knowledge which are reported in this study in section 10.2. The implications for academia and health organisations are presented in section 10.3 as well as a reflection on the ethical challenges of this study. The following section (10.4) discusses the limitations of the study. The chapter concludes with recommendations for future work (section 10.5) and some final remarks (section 10.6).
10.1 Summary of findings

This section is a descriptive account of the results, and this summary is developed from the findings of each results chapter. The results from Chapter 6 concerned an approach to apply the Twitter data beyond the tweet to contextualise the description of the sample and reveal inequalities across gender, age, location and type of place of the young people in contact with The Filter Wales. The total number of sample members was 2180, 56.6% of which were women. Only 27% of the sample members had an age indication and demonstrated that they were between 15 and 36 years old, but the vast majority were aged 18 to 22. The online smoking status was gathered through the smoking-related tweets of each individual in the sample, and the analysis estimated that there were mostly non-smokers in the sample (38.4%), followed by smokers (31.8%), quitters (18.5%) and lastly relapers (11.3%).

The likely place of residence was determined for 1051 (48.2%) of the sample population by extracting the geolocation that was attached to their tweets. This revealed that the sample lived across the whole of Wales but mainly in the more densely populated south. These places of residence for people could further be examined by deprivation level and rural/urban classification. The sample members resided across all quintiles of deprivation to an almost equal measure. There was some variability with smoking status as quitters were more present in less advantaged neighbourhoods and a higher proportion of non-smokers was found in the affluent areas. The variation in rural/urban residence was minimal for gender and smoking status, and almost 78.8% of the sample population lived in an urban area. The sample population is not a representation of young people in Wales but provides insights in that The Filter Wales has reached people from all over Wales regardless of where their outreach events have taken place (most are Cardiff centred). The results also demonstrate that the campaign is reaching out to young people across the whole range of areas according to deprivation status.

In Chapter 7, the quantitative content analysis of the (16,688) smoking-related tweets identified the smoking product (i.e. tobacco, marijuana, e-cigarettes, and shisha), the person performing the action (the originator or someone else), and the smoking activity (i.e. smoking, desire to smoke, thinking of quitting and quitting smoking) referenced in each tweet. This gave a clear overview of what young people found most important smoking content to tweet. The majority of the tweets were made on the topic of tobacco and concerned the originator of the tweet smoking tobacco. Marijuana was the second most tweeted about smoking product and the only product in this study that showed an apparent
gender difference in the proportion of tweets relating to marijuana and the type of content concerning this product. Men tweeted more on this topic and were more likely to tweet about smoking marijuana themselves whereas women tweeted more about others smoking marijuana. The assumed age variable showed the important find that even in the youngest age group individuals tweeted about serious quitting smoking attempts.

The temporal qualities of the data were explored through plotting the smoking-related tweets (from July 2013 to June 2016) onto time charts of hours of the day, days of the week, and months of the year. These times refer to the times and dates the tweets were posted, not necessarily when the smoking products were consumed. The results showed that tobacco and e-cigarettes are activities that were tweeted about at any given time, but that marijuana and shisha smoking tweets were mostly posted at ‘social times’ in the evening and during weekdays. The longer timeline of the weeks in a year showing only tobacco-related tweets illustrated that the sample members tweeted about different tobacco activities and topics (besides the smoking of tobacco) such as the plain packaging regulation in May 2015.

Sentiment analysis was applied to the content of the tweets, and the sentiment scores that were derived revealed that the sample members tweeted most positively about quitting smoking and shisha and tweeted most negatively about marijuana. After placing the average sentiment scores on timelines, the young people tweeted more negatively than positively about tobacco, but this fluctuated for the other smoking products.

The qualitative content analysis of the health co-behaviours in the smoking-related tweets, in Chapter 8, demonstrated that gender was important in the tweeting of this content; women tweet more about (un)healthy co-behaviours and men about smoking and alcohol. The smoking content of health co-behaviour tweets concluded that young people know what is unhealthy and what is healthy but that most of the time the unhealthy choices prevail over the healthy ones. This resulted in a fair number of tweets containing references to smoking while trying to be healthy and how having a cigarette is obstructive to their health.

The last section of the results from Chapter 9 summarises the tweet context of the smoking tweets through the analysis of discourse across the entire Twitter feeds of a subsample of fifty randomly selected sample members. First, the Twitter style analysis revealed that smokers tweet more emotionally and quitters have a more positive online persona. Going further into the online context of the smoking tweets, the tweets are created to induce a response or inform others about their life, with tweets reporting ‘having quit smoking’
being particularly popular. The Twitter context of the tweets examined across this sample mostly connected to motives for using Twitter (i.e. general interest and socialising), the Twitter Style of these subsample members, and the way they interact with like-minded people on Twitter.

10.2 Contributions to knowledge

This section reports the central contributions to knowledge that this thesis provides. Where relevant, this section also outlines how these relate to debates within existing literature. Bearing in mind the aims of the study this section focuses on the contribution this study makes towards the evaluation of The Filter Wales and the understanding of the social meaning and context of smoking. In this section, the results are taken from their attributing chapters and are placed on a broader framework of how these results contribute to knowledge of youth smoking.

10.2.1 The reach of the Twitter element of The Filter Wales

The first contribution is that this study combines the efforts of The Filter Wales with the personal insights of youth to better understand youth smoking and when it is best to engage these young people in anti-tobacco interventions. The reach of the smoking cessation organisation can be explained in various ways, and the following paragraphs outline reach regarding the description of the people The Filter Wales campaign engaged with on Twitter.

Meek, Hurt & Grant’s evaluation of The Filter Wales (2015) was unable to establish much about the young people that the Filter Twitter element had been in contact with as the Twitter handle was only a small part of their evaluation. Only eight young people were interviewed, and 21 responded to the online survey to express their opinions on the social media handle of The Filter Wales, not express their characteristics or understanding of smoking. The study presented here uncovered gender, (for a small group) age, Twitter activity, smoking status, and (where possible) place of residence of 2180 young people in Wales which indicated that the Filter engaged with their target group (i.e. both men and women mainly between the age of 11 and 25 from all across Wales), an observation which was previously unknown. While Meek et al. (2015) laid the groundwork for an assessment of the reach of the Twitter element of the Filter; this study provides a more extensive examination by covering all of the young people that have been in contact with the Filter. This thesis also provides a methodological contribution to the analysis of the campaign’s reach through mapping the tweets’ geolocation. The method applied here was not
undertaken in the previous evaluation by Meek et al. and has not been used by the
organisers of the Filter campaign. Acquiring the Twitter archive of different individuals
provided the possibility to locate where a person lived which created an output that was
larger than would be expected with traditional data collection methods; on average only 2-5% of all tweets being coupled with geolocation coordinates (Burton et al., 2012; Leetaru et al., 2013). This possibility is underused especially in regards to adopting a place of
residence for social research into Twitter users. Other health studies have based the
location purely on from where an individual the tweet was sent (e.g. Ghosh & Guha, 2013;
Graham, Hale, & Gaffney, 2014) or the location is derived from the personal information
on their Twitter profiles (e.g. Murnane & Counts, 2014).

Twitter presented the opportunity to uncover the ‘geolocation’ of the tweets made by the
young people, and with that, an indication of the likely place of residence for 1051 sample
members could be made. The use of the Twitter archive and only taking the people that
have multiple geolocated tweets establishes the likely place of residence better than solely
relying on the coordinates of single geolocated tweets. Linking the place of residence to
characteristics of the place (in this case deprivation levels and rural/urban classification)
contributes to the health and place literature as these finding presented here are
contextualised across types of places and not merely geographical location. This study
contributes to the ‘geography’ of smoking by signifying where the Twitter users live
through multiple inputs and illustrate the social inequalities based on locational
characteristics.

To have an effective anti-smoking campaign, it is critical to uncover if the online platform
of choice is a good way to reach young people. The exploration of the Twitter activity and
Twitter style showed that young people are already using Twitter extensively as was
shown by other studies (e.g. Duggan & Brenner, 2013; Lenhart, Purcell, Smith, & Zickuhr,
2014). Moreover, the Twitter archives of individuals with extensive information about the
young people’s online personas allowed for assumed online smoking status and this
revealed that The Filter has been in contact with and reached different young people.
Information on youth smoking status is generally derived from questionnaires about
smoking habits which can be associated with other variables (such as in the HSBC survey).
In this study, smoking status was not derived from direct questioning. Instead, through
necessity, it was derived from an analysis of the text contained in the sample members’
tweets. Although there is little opportunity to verify the assumed smoking status
constructed in this way, its derivation, which is less intrusive than direct face-to-face
questioning may well improve the validity of the information. This validity question has been raised previously in the literature review as young people were unlikely to perceive themselves as smokers (e.g. Amos et al., 2006; Berg et al., 2009; Heikkinen et al., 2010; Leas, Zablocki, Edland, & Al-Delaimy, 2014). Moreover, this way of assessing smoking status avoids recollection bias about smoking behaviour that was prompted by Mair et al. (2006) questioning people about smoking initiation with responses that often changed with each consecutive year or by Berg et al. (2009) on their study on smoking cessation and how people forgot that they made a quitting smoking attempt.

The quantitative content analysis of Chapter 7 aided in this question of reach by examining what the young people in the sample tweeted about smoking. Many tweets on the topic of smoking were made besides the ones picked up by The Filter Wales. So, regarding reach, The Filter Wales has reached people that posted more on smoking than purely the tweet the organisation found. However, the organisation has possibly only reached young people that are more engaged with smoking, to begin with, as it was important enough for them to tweet about smoking unprovoked by the Filter. The contribution to a better grasp of the reach of The Filter Wales showed that the campaign has engaged with the target group from all across Wales but that these young people already found smoking interesting enough to tweet about which gives no indication that these are hard-to-reach or ‘at risk’-youths.

10.2.2 A better understanding of the social meaning of smoking and health risk co-behaviours

This thesis incorporates novel and traditional methods to Twitter data to provide a better understanding of smoking-related tweets and the young people who made them. By using Twitter data (which is an unprovoked output of young people’s social interactions), the results are quite possibly a better representation of the meaning they give to smoking and health risk co-behaviours than would be uncovered through direct questioning.

Confirmed by surveys as the most consumed out of the different smoking products (ASH, 2015c; Fuller et al., 2014; HBSC, 2015), tobacco smoking seems well-established in the everyday lives of the young people. One of the contributions of this study is that tobacco is so well established that many of the tobacco tweets do not directly reference tobacco (but just smoking) and that a fair amount of the tobacco tweets are on an entirely different topic. This result challenges the importance of tobacco smoking in young peoples’ lives. Tobacco smoking is only mentioned as an activity while something else happened which
would not be uncovered if the sample members were directly questioned about their smoking behaviour.

There is only a small base of literature (mainly questionnaires and surveys such as the HBSC and the SDD) that contrasts different smoking products, and understanding what drives the perceptions of each of them is largely underdeveloped. Studies suggest that electronic cigarettes are now more popular than tobacco and is considered a type of smoking in its own right (De Lacy et al., 2017; Goldstone et al., 2016; Wang, Wang, Cao, Wang, & Hu, 2016). This was not reflected in the analysis of the Twitter feeds of the sample. The tweets found in the content analyses and the analysis of discourse of e-cigarette smoking all refer to quitting tobacco smoking. Shisha smoking is another emerging product, and the majority of the literature suggests that its use is rising in the UK (Akl et al., 2015). This study contributes to this literature by showing how shisha smoking is, on the whole, not considered a smoking product and more commonly seen as part of a pleasant social activity.

Surveys often examine what individual health behaviours any one person engages in, but they rarely question how they interplay or interact with each other. Previous literature (e.g. Mistry, McCarthy, Yancey, Lu, & Patel, 2009; Peters et al., 2009; van Lenthe et al., 2009; Wiefferink et al., 2006) has shown that engagement in health risk behaviours tends to cluster; if people smoke, they are also more likely to consume alcohol, have an unhealthy diet, and exercise below the recommended guidelines. This thesis further develops these notions of clustering and shows not only how these health behaviours are connected but also how these interactions are different depending on the smoking activity and health behaviour. This study has shown, for example, how especially physical activity suffers when an individual is also a smoker and that several individuals focus on contradictions in their own behaviour (e.g. going for a run and having a cigarette afterwards or having quit smoking but relapse when they are consuming alcohol).

The content analysis and analysis of discourse of the tweets in this study revealed marginalisation towards smokers. Chapter 7 showed 1954 (18.8%) tweets about ‘others’ smoking, and the overall positive sentiment score was lower for these tweets. These tweets indicate that judgement towards smokers is present and apparently important enough to post this Twitter content online. This marginalisation was already uncovered through traditional methods in many studies (e.g. Alexander, Frohlich, Poland, Haines, & Maule, 2010; Frohlich, Mykhalovskiy, Poland, Haines-Saah, & Johnson, 2012; Parry, Carnwell, Moore, & Murphy, 2010; Ritchie, Amos, & Martin, 2010). The findings of this study
contribute to the literature on smoking marginalisation by showcasing how a space such as Twitter can intensify the sense of marginalisation as anybody can post anything about anyone. This is more clearly exemplified in Chapter 9 in which some young people tweet marginalising content about others (e.g. judgements on smoking pregnant women). These tweets are not directed at the smokers in question as they are not likely to know or follow each other on Twitter. The danger of repercussion of the originator is low which makes it likely that the tweet is created to interest their followers that may tweet similar content. The sense of anonymity and the gain of bonding with like-minded people on Twitter increase the idea that judging others is acceptable.

Even more concerning is the self-stigmatisation present in the tweets. The aversion of smoking has made the smokers aware of their own behaviour, and when combined with the difficulty of quitting can lead to serious (mental) health issues (Evans-Polce, Castaldelli-Maia, Schomerus, & Evans-Lacko, 2015). To the researcher’s knowledge, this self-stigmatisation in Twitter data has not been revealed and discussed before. This finding signals towards the overall divergence of knowledge about health issues, not having the resources to change them and experiencing marginalisation from an anonymous group through social network sites.

A central theme throughout the results chapters is to better understand the meaning of smoking by comparing the smoking tweets to general Twitter activity. This is challenging the methods of many health studies (e.g. Jamison-Powell, Linehan, Daley, Garbett, & Lawson, 2012; Sullivan et al., 2012; West et al., 2012) which all looked at a specific health topic and uncovered its occurrence and (occasionally) content by collecting only the tweets on that topic. Chapter 7 exemplified the divergence by adding a ‘general Twitter activity’ timeline to the time charts to illustrate how the smoking-related tweets are substantially different from general ones. The Twitter data revealed when young people tweeted about smoking and not when they are actually smoking, indicating that smoking is only mentioned when it is relevant for them to mention it such as tweeting about new anti-tobacco regulation. These smoking-related tweets represent something more than mere Twitter activity and are posted independently of when these young people usually tweet. The contribution of showcasing this divergence is something that was expected but not researched in the case of smoking.
10.3 Implications

As discussed above, this study makes several contributions to the emerging literature on meaning and context of youth smoking. These contributions have implications for academia and health organisations and reveal some specific ethical challenges. By understanding the complex meaning and interpretation young people give to smoking, a start can be made to align the anti-tobacco initiatives to prevent smoking uptake better and more efficiently influence youth smoking cessation.

10.3.1 Academic implications

The use of Twitter as a research tool has been growing in social research in the last years and is likely to expand even further. The overall implication for academia is that Twitter is low-cost, fast and provides voluminous data. It is an almost endless source of big data for social science that is as yet underused. Twitter provides information that would be difficult to uncover through face-to-face interactions and can be used to uncover more deeply rooted perceptions.

Most smoking tweets seem random in their Twitter histories; however, when analysing them further for indications of discourse, patterns emerge of when these smoking-related tweets are relevant, and a narrative can be found by using the Twitter history of a person, not just the tweet containing the topic of the study as mentioned before. This has been shown to provide a more complete picture of the person posting the tweets and provides valuable indications for the examination of the topic at hand. With the assistance of linguistics programs, this information can be placed in a broader framework of how young people tweet about smoking by taking into consideration how they use Twitter. Most importantly, applying traditional methods such as content analysis and analysis of discourse on Twitter data added a new layer to smoking and Twitter research.

As a consequence of this research, academics can use the methods of determining place of residence through multiple geolocation entries. This information provides more reliable data on where the person most frequents instead of the specific location from the tweet was sent. This place of residence has more advantages as the way in which it is applied in this study; the individuals merge into aggregates of people which provides a form of anonymity to the individuals and the aggregate level analysis (i.e. lower super output areas) allowed for added examination of the characteristics of the location.

The last implication for smoking research presented here is that the people aged 11 to 25 form a standalone group that is particularly relevant for smoking research as 99% of the
smokers started within this age range (ASH, 2015a). Too often in smoking surveys, the cutoff point is 16-years-old (see for example Statistics of Wales Health Survey) which breaks up this group that together form the key to the smoking endgame.

10.3.2 Implications for health organisations

The results of this thesis show that the Filter’s decision to contact people via Twitter substantially improves the reach of the campaign beyond the contacts made through traditional outreach activities. This finding has implications for many other health organisations who need to engage with young people as an online approach may be more successful than community outreach for health messaging and campaign engagement. The findings and methods of this study can be replicated and applied to other health organisations taken that they are actively searching for their target audience. Future work should consist of having more organisations contact their target audience in this way so that the outreach of health information and resources expand to a broader audience who would not actively search for this information themselves. It comes at little cost and can increase awareness considerably as shown by the results of this thesis.

Smoking cessation organisations (and anti-tobacco control) focus their efforts on tobacco smoking and the assistance of smokers wanting to quit tobacco as it is still the most commonly used smoking product. This is confirmed by the results of this study. As a consequence of this study, The Filter Wales might increase the focus on the effects of e-cigarettes and shisha smoking in combination with encouraging stronger regulations for the consumption of these products before they become normalised and accepted by society. With the more subtle taste, these products are more appealing to young people (Myslín et al., 2013), and more attention should be focused on preventing the uptake of these products, even if young people only smoke these sporadically.

As suggested in Chapter 9, the Twitter style varied between individuals with different smoking status, e.g. quitters were more positive, and smokers were more emotional. To engage with these people, the exploration of Twitter style might be convenient as people like to engage with people who are similar (Marwick & Boyd, 2010; Rosenquist et al., 2011; M. S. Smith & Giraud-Carrier, 2010). From reading the communication by The Filter Wales, this has already been done as the social media team responds with a similar style to the original tweet, but other youth health organisations might want to consider this approach for their online outreach campaigns.
Young people are well aware of what is healthy and not when it comes to health behaviours as shown in Chapter 8, but the normalisation of smoking, particularly in connection with alcohol use, persists. As the results from this thesis exemplify; the awareness of the possible health problems are there, but the unhealthy choice is often too tempting, and young people have difficulty keeping to a healthy (non-smoking) lifestyle. This is further confirmed by the detection of self-stigmatisation which has implications for health organisations and tobacco control as it shows an adverse consequence of their efforts. The self-stigmatising indicates that smokers are very much aware of their damaging behaviour but internalise the marginalisation instead of changing their habit. The implication being that health organisations need to refocus on helping people change co-behaviours in addition to providing judgement-free information.

10.3.3 Reflection on ethical challenges

The ethical challenges of working with Twitter data are that the people that post on Twitter are not always aware that their data can and is used by third parties (Mao et al., 2011; McKee, 2013; Rivers & Lewis, 2014). The debate here concerns if it is ethical to use someone’s data in the study if the individual is unaware of their contribution and to what extent it can cause harm. Kelley & Cranshaw (2013) call this ‘user privacy expectations’ indicating that people posting on social network sites do that for no other reason than to interact with others online. Another issue arises as individuals are generally unable to assess the implications of interacting on social network sites and provide personal information that might threaten their privacy (Solove, 2013). Mao et al. (2011) studied privacy leaks on Twitter and showed that a significant portion of tweets contained personal information. In addition, the extra information (such as geolocation and specific time and date) makes it easy for third parties to pinpoint where someone is at that specific time. This study has addressed these challenges by using aggregates of people for most of the analysis and anonymising the sample members in other parts. Furthermore, place of residence is described via a geographical area (i.e. the lower super output area) and not specific coordinates. It would be good practice if this approach was applied more often in research bodies in academia or industry.

These ethical debates will continue as these ethical challenges remain only guidelines and not hard regulation. The lack of regulations can be particularly damaging to individuals who are unaware of the consequences of posting on Twitter. There is evidence, however, to suggest that the use of Twitter is changing into a more information-based platform which has less personal content and this could reduce some of the ethical challenges. Also,
the majority of young people have now moved on to other social network platforms such as Snapchat and Instagram (Duggan, 2015) and ethical challenges have shifted to these alternative social network platforms that do not offer data to be used by third parties.

10.4 Study limitations

The research presented in this thesis contributes to the growing body of work on the use of social media in public health research, providing an important contribution for academic and policy audiences concerned with smoking. However, the findings stated should be considered in conjunction with some limitations.

10.4.1 Limitations of the study sample

One of the limitations concerns the guarantee that all the sample members are young people from the Wales area. Anyone can create a Twitter profile and fill it with false information. However, due to the thorough study of the content of the tweets and profiles of each individual, it can be assumed that they are real people. Related to this, it is uncertain if the sample members are all young. The age indication was only present for 27% of the sample members, and there was no verification that this age indication was accurate. Similar difficulties are found in the place of residence indicator. The geolocation coordinates were found for 52% of the sample population, and the analysis could give a highly biased estimation of the place of residence. All these exclusions caused a non-representative sample, and all findings are limited to the people in contact with The Filter Twitter account.

This study did not interact with the young people in the sample as this might change the way they behave in the Twitter feeds if these were collected after this interaction. However, the data might have changed after The Filter Wales engaged with these individuals on Twitter. Receiving a message from an account outside of the Twitter network provides an instant reminder to the individuals that their data is found by different people than they (might have) intended and might change the way they tweet (particularly about smoking). By cooperating with this smoking cessation organisation, this limitation could not be avoided.

The data and the users that were selected for this study are based on the Twitter activity of the Filter, and there are some limitations to using the Filter Twitter as data. The Filter social media team were selective in their contact with users as not all smoking-related tweets they found were replied to or retweeted. Notably, The Filter Wales did not want to associate themselves with the content of the tweet when the tweet was discriminative or
exceedingly rude. Furthermore, out of all the people in Wales that tweeted about smoking in the defined time period, only an undefined proportion was picked up and interacted with by The Filter Wales.

10.4.2 Limitations of Twitter data compared to traditional methods

The limitations of Twitter compared to traditional methods show that Twitter data is abundant, but that the information is driven by the participants which led to many collected tweets not being suitable for the study. Likewise, the Twitter archive does not form a cohesive structure, and the Twitter archive data do not contain replies from other Twitter users, so the data can only be examined from the side of the sample members’ output without knowing what happened in between those tweets.

Twitter data is generated with an exact time and date, but it can, therefore, pick up incidences that happened once. For example, someone can have one cigarette and tweet about smoking it which would label (for the purposes of analyses in this study) the individual as being a smoker. However, this act of smoking may have been an isolated event, and in this sense, while using Twitter data avoids recall bias, it does not necessarily reflect normal (i.e. dominant), behaviour and perceptions. Similarly, the Twitter activity might be a cause for determining smoking status. People who use Twitter to inform others are more likely only to tweet success stories such as having quit smoking. Equally, when individuals use Twitter as an online diary, then the tweets become more personal and are more likely to showcase everyday events such as smoking.

A significant challenge came from using Sentistrength for sentiment analysis. This automated program was designed for microblogs specifically but was not well adjusted for the writing style of the young people in the sample. Twitter has limited space for text (140 characters), and therefore many words are changed, abbreviated, or distorted (e.g. ‘because’ becomes ‘cos’) and those words are not picked up by these programs for their intended meaning (Sumner, Byers, Boochever, & Park, 2012). Moreover, a limitation to these sentiment analyses is that it does not account for jokes, urban slang, or sarcasm (Sumner et al., 2012; Thelwall et al., 2010) which is a particular concern when looking at tweets from this group. This limitation led to some tweets being scored differently from their intended meaning as young people used phrases that could mean something different in academia than for these young people and their followers, e.g. ‘sick’ is considered positive among youth.
10.5 Recommendations for future work

This research provided the groundwork for many possible avenues of future work, and this section presents three main areas. As mentioned in the limitations, the sample population was provided by The Filter Wales who were selective in their engagement with young Twitter users. For future work, the middleman (the Filter) could be removed to get a full picture of the tweets on the topic of smoking in a particular area whereby the researchers set up their own search engine and gather all tweets regardless of the content and continue with the step taken in this study. This would comprise a more extensive sample of all the tweets that can be analysed within the chosen parameters.

This study was intentionally designed to only focus on the information available from the Twitter accounts from the sample members without pursuing interaction or additional information. For future work, it might be beneficial to contact the people in the study and gain additional ‘reliable’ information about the sample members regarding, for example, their age, place of residence and motivation for a smoking cessation attempt. This recommendation does come with added ethical challenges as respondents tweet primarily to interact with others and may not approve (when confronted) of their information being used by third parties in this way.

The focus of this study was smoking, and all the Twitter content used in this study contained a reference to smoking. Chapter 8 has extended the content analysis to health co-behaviours (with smoking). This study could easily be extended to other health behaviour research provided that the group has been in contact with a health organisation.

10.6 Final remarks

This thesis was written to critically examine the way in which young people, contacted through The Filter Wales Twitter platform, give meaning to smoking in the text of their tweets. In doing so, this research has uncovered that the awareness of the health risks of this deadly habit is high. However, thus far, the positive associations with smoking such as smoking being de-stressing and a good way to socialise are juxtaposed against these integrated messages. The insights supplied by this thesis present the opportunity for anti-smoking organisations to increase their influence by adding a social media aspect to their programme and has shown that with some encouragement a healthier lifestyle for young people could be possible.
Bibliography


Cobb, C. O., Shihadeh, A., Weaver, M. F., & Eissenberg, T. (2011). Waterpipe tobacco smoking and


Culotta, A. (2010). Towards detecting influenza epidemics by analyzing Twitter messages. *1st Workshop on Social Media Analytics.*


221


Johnson, P. R., & Yang, S.-U. (2009). Uses and Gratifications of Twitter. In *the Communication Technology Division of the annual convention of the Association for Education in Journalism and Mass Communication* (pp. 1–32). Boston, Massachusetts.


Moore, L., Roberts, C., & Tudor-Smith, C. (2001). School smoking policies and smoking prevalence among


comparison of three recruitment mechanisms. *Nicotine & Tobacco Research, 12*(7), 768–75.
https://doi.org/10.1093/nttr/ntq086


https://doi.org/10.1207/s15324796abm3003_9


Struik, L. L., & Baskerville, N. B. (2014). The role of Facebook in Crush the Crave, a mobile- and social
media-based smoking cessation intervention: qualitative framework analysis of posts. *Journal of Medical Internet Research, 16*(7), e170. https://doi.org/10.2196/jmir.3189


236

Appendices

Appendix A – Ethics committee review letter
  – Ethical conduct declaration form
Appendix B – SQL*Developer query
Appendix C – LSOA validity check table
Appendix D – Health risk co-behaviour codes
Appendix E – Table of the description of 50 subsample members
Appendix F – Codes in Atlas.ti
Appendix G – Twitter Style analysis tables
Appendix A – Ethics committee review letter

Science Faculty Ethics Committee
Faculty of Science
University of Portsmouth

Department of Psychology
University of Portsmouth
cornelia.van-diepen@port.ac.uk

5th January 2016

FAVOURABLE ETHICAL OPINION

Study Title: The use of Twitter to support smoking-related public health activities surrounding adolescents in Wales.
Reference Number: SFEC 2015-103 (Please quote this in any correspondence)

Thank you for submitting your application to the Science Faculty Ethics Committee (SFEC) dated 15/12/15 in accordance with current procedures.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions.

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

Yours sincerely,

[Signature]

Dr Chris Markham
Chair, Science Faculty Ethics Committee

1 Procedures for Ethical Review. Science Faculty Ethics Committee, University of Portsmouth, October 2012 (to be updated).
2 After ethical review – Guidance for researchers (Please read).
Appendix A (continued) – Ethical conduct declaration form

**FORM UPR16**
Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Postgraduate Research Student Handbook for more information)

<table>
<thead>
<tr>
<th>Postgraduate Research Student (PGRS) Information</th>
<th>Student ID: UP788844</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGRS Name: Cornelia van Dijmen</td>
<td></td>
</tr>
<tr>
<td>Department: Geography</td>
<td></td>
</tr>
<tr>
<td>First Supervisor: Prof. Liz Twigg</td>
<td></td>
</tr>
<tr>
<td>Start Date: February 2015</td>
<td></td>
</tr>
<tr>
<td>(for progression date for Prof Doc students)</td>
<td></td>
</tr>
<tr>
<td>Study Mode and Route:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part-time</td>
</tr>
<tr>
<td>Title of Thesis: A Twitter-based study on the reach of a smoking cessation organisation and the social meaning of smoking</td>
<td></td>
</tr>
<tr>
<td>Thesis Word Count: (excluding ancillary data)</td>
<td>72747</td>
</tr>
</tbody>
</table>

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University’s Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study.

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-fund/etds-of-practice-for-research/)

<table>
<thead>
<tr>
<th></th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>b) Have all contributions to knowledge been acknowledged?</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>c) Have you complied with all agreements relating to intellectual property, publication and authorship?</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>e) Does your research comply with all legal, ethical, and contractual requirements?</td>
<td>x</td>
<td>✓</td>
</tr>
</tbody>
</table>

Candidate Statement:
I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC): SFEC:2015-103

If you have not submitted your work for ethical review, and/or you have answered ‘No’ to one or more of questions a) to e), please explain below why this is so:

Signature (PGRS): [Signature]

Date: 30-01-2020
Appendix B – SQL*Developer query

```sql
select * FROM TOTALTWEETS
WHERE (content like '%fag%')
  or (content like '%cig%')
  or (content like '%baccy%')
  or (content like '%roppy%')
  or (content like '%weed%')
  or (content like '%tobacco%')
  or (content like '%smoking%')
  or (content like '%stoptober%')
  or (content like '%thefilter%')
  or (content like '%cannabis%')
  or (content like '%splif%')
  or (content like '%vapour%')
  or (content like '%stoned%')
  or (content like '%marlboro%')
  or (content like '%waterpipe%')
  or (content like '%marijuana%')
  or (content like '%hooka%')
  or (content like '%shisha%')
```
Appendix C – LSOA validity check Table

The validity check resulted in the same LSOA for 89.24% of the sample members with a place of residence indication in Wales.

<table>
<thead>
<tr>
<th>Username</th>
<th>LSOA from first number</th>
<th>LSOA after validity check</th>
<th>Same LSOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person1</td>
<td>Halesowen Gwent 001E</td>
<td>Halesowen Gwent 001E</td>
<td>yes</td>
</tr>
<tr>
<td>Person2</td>
<td>Halesowen Gwent 002B</td>
<td>7</td>
<td>no</td>
</tr>
<tr>
<td>Person3</td>
<td>Halesowen Gwent 004B</td>
<td>321</td>
<td>199 yes</td>
</tr>
<tr>
<td>Person4</td>
<td>Halesowen Gwent 007C</td>
<td>47</td>
<td>15 yes</td>
</tr>
<tr>
<td>Person5</td>
<td>Halesowen Gwent 007E</td>
<td>90</td>
<td>60 yes</td>
</tr>
<tr>
<td>Person6</td>
<td>Halesowen Gwent 008C</td>
<td>158</td>
<td>79 yes</td>
</tr>
<tr>
<td>Person7</td>
<td>Halesowen Gwent 008F</td>
<td>1766</td>
<td>1252 yes</td>
</tr>
<tr>
<td>Person8</td>
<td>Bridgend 001A</td>
<td>45</td>
<td>27 yes</td>
</tr>
<tr>
<td>Person9</td>
<td>Bridgend 001C</td>
<td>3</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Bridgend 001C</td>
<td>373</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Bridgend 002C</td>
<td>56</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Bridgend 002C</td>
<td>54</td>
<td>36 no</td>
</tr>
<tr>
<td>Person13</td>
<td>Bridgend 002C</td>
<td>71</td>
<td>55 yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Bridgend 002C</td>
<td>1545</td>
<td>988 yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Bridgend 002E</td>
<td>1300</td>
<td>859 yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Bridgend 006A</td>
<td>1374</td>
<td>918 yes</td>
</tr>
<tr>
<td>Person17</td>
<td>Bridgend 006A</td>
<td>168</td>
<td>136 yes</td>
</tr>
<tr>
<td>Person18</td>
<td>Bridgend 006A</td>
<td>1011</td>
<td>546 yes</td>
</tr>
<tr>
<td>Person19</td>
<td>Bridgend 006C</td>
<td>10</td>
<td>3 yes</td>
</tr>
<tr>
<td>Person20</td>
<td>Bridgend 007D</td>
<td>62</td>
<td>46 yes</td>
</tr>
<tr>
<td>Person21</td>
<td>Bridgend 007D</td>
<td>35</td>
<td>26 yes</td>
</tr>
<tr>
<td>Person22</td>
<td>Bridgend 008A</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>Person23</td>
<td>Bridgend 008A</td>
<td>9</td>
<td>yes</td>
</tr>
<tr>
<td>Person24</td>
<td>Bridgend 009A</td>
<td>23</td>
<td>yes</td>
</tr>
<tr>
<td>Person25</td>
<td>Bridgend 009B</td>
<td>9</td>
<td>yes</td>
</tr>
<tr>
<td>Person26</td>
<td>Bridgend 009B</td>
<td>225</td>
<td>yes</td>
</tr>
<tr>
<td>Person27</td>
<td>Bridgend 009B</td>
<td>27</td>
<td>yes</td>
</tr>
<tr>
<td>Person28</td>
<td>Bridgend 009C</td>
<td>153</td>
<td>yes</td>
</tr>
<tr>
<td>Person29</td>
<td>Bridgend 009C</td>
<td>128</td>
<td>yes</td>
</tr>
<tr>
<td>Person30</td>
<td>Bridgend 009E</td>
<td>4</td>
<td>yes</td>
</tr>
<tr>
<td>Person31</td>
<td>Bridgend 009E</td>
<td>117</td>
<td>yes</td>
</tr>
<tr>
<td>Person32</td>
<td>Bridgend 010B</td>
<td>246</td>
<td>yes</td>
</tr>
<tr>
<td>Person33</td>
<td>Bridgend 010D</td>
<td>52</td>
<td>yes</td>
</tr>
<tr>
<td>Person34</td>
<td>Bridgend 011A</td>
<td>129</td>
<td>yes</td>
</tr>
<tr>
<td>Person35</td>
<td>Rhondda Cynon Taf 012A</td>
<td>29</td>
<td>yes</td>
</tr>
<tr>
<td>Person36</td>
<td>Rhondda Cynon Taf 012B</td>
<td>21</td>
<td>yes</td>
</tr>
<tr>
<td>Person37</td>
<td>Rhondda Cynon Taf 012B</td>
<td>784</td>
<td>yes</td>
</tr>
<tr>
<td>Person38</td>
<td>Rhondda Cynon Taf 012B</td>
<td>565</td>
<td>yes</td>
</tr>
<tr>
<td>Person39</td>
<td>Bridgend 011C</td>
<td>30</td>
<td>yes</td>
</tr>
<tr>
<td>Person40</td>
<td>Bridgend 011C</td>
<td>26</td>
<td>yes</td>
</tr>
<tr>
<td>Person41</td>
<td>Bridgend 011C</td>
<td>2335</td>
<td>yes</td>
</tr>
<tr>
<td>Person42</td>
<td>Bridgend 012D</td>
<td>333</td>
<td>yes</td>
</tr>
<tr>
<td>Person43</td>
<td>Bridgend 012D</td>
<td>278</td>
<td>yes</td>
</tr>
<tr>
<td>Person44</td>
<td>Bridgend 013B</td>
<td>6</td>
<td>yes</td>
</tr>
<tr>
<td>Person45</td>
<td>Bridgend 013B</td>
<td>1168</td>
<td>yes</td>
</tr>
<tr>
<td>Person46</td>
<td>Bridgend 013C</td>
<td>943</td>
<td>yes</td>
</tr>
<tr>
<td>Person47</td>
<td>Bridgend 013C</td>
<td>11</td>
<td>yes</td>
</tr>
<tr>
<td>Person48</td>
<td>Bridgend 014A</td>
<td>29</td>
<td>yes</td>
</tr>
<tr>
<td>Person49</td>
<td>Bridgend 014A</td>
<td>54</td>
<td>yes</td>
</tr>
<tr>
<td>Person50</td>
<td>Bridgend 014A</td>
<td>9</td>
<td>yes</td>
</tr>
<tr>
<td>Person51</td>
<td>Bridgend 014A</td>
<td>21</td>
<td>yes</td>
</tr>
<tr>
<td>Person52</td>
<td>Bridgend 014A</td>
<td>11</td>
<td>yes</td>
</tr>
<tr>
<td>Person53</td>
<td>Bridgend 014A</td>
<td>95</td>
<td>yes</td>
</tr>
<tr>
<td>Person54</td>
<td>Bridgend 014C</td>
<td>217</td>
<td>yes</td>
</tr>
<tr>
<td>Person55</td>
<td>Bridgend 014C</td>
<td>643</td>
<td>yes</td>
</tr>
<tr>
<td>Person56</td>
<td>Bridgend 015A</td>
<td>327</td>
<td>yes</td>
</tr>
<tr>
<td>Person57</td>
<td>Bridgend 015A</td>
<td>219</td>
<td>yes</td>
</tr>
<tr>
<td>Person58</td>
<td>Bridgend 015D</td>
<td>169</td>
<td>yes</td>
</tr>
<tr>
<td>Person59</td>
<td>Bridgend 016A</td>
<td>7</td>
<td>yes</td>
</tr>
<tr>
<td>Person60</td>
<td>Bridgend 016A</td>
<td>761</td>
<td>yes</td>
</tr>
<tr>
<td>Person61</td>
<td>Bridgend 017D</td>
<td>11</td>
<td>yes</td>
</tr>
<tr>
<td>Person62</td>
<td>Bridgend 017E</td>
<td>743</td>
<td>yes</td>
</tr>
<tr>
<td>Person63</td>
<td>Bridgend 017F</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>Person64</td>
<td>Bridgend 017F</td>
<td>42</td>
<td>yes</td>
</tr>
<tr>
<td>Person65</td>
<td>Bridgend 017G</td>
<td>17</td>
<td>yes</td>
</tr>
<tr>
<td>Person66</td>
<td>Bridgend 018A</td>
<td>470</td>
<td>yes</td>
</tr>
<tr>
<td>Person67</td>
<td>Bridgend 018B</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>Person68</td>
<td>Bridgend 018C</td>
<td>1336</td>
<td>yes</td>
</tr>
<tr>
<td>Person69</td>
<td>Bridgend 018C</td>
<td>46</td>
<td>yes</td>
</tr>
<tr>
<td>Person70</td>
<td>Bridgend 018C</td>
<td>1531</td>
<td>yes</td>
</tr>
<tr>
<td>Person71</td>
<td>Bridgend 018D</td>
<td>170</td>
<td>yes</td>
</tr>
<tr>
<td>Person72</td>
<td>Bridgend 018E</td>
<td>1256</td>
<td>yes</td>
</tr>
<tr>
<td>Person73</td>
<td>Bridgend 018E</td>
<td>507</td>
<td>yes</td>
</tr>
<tr>
<td>Person74</td>
<td>Bridgend 019B</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>Person75</td>
<td>Bromsgrove 003C</td>
<td>78</td>
<td>yes</td>
</tr>
<tr>
<td>Person</td>
<td>Address</td>
<td>Area</td>
<td>Yes/No</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>Person67</td>
<td>Cardiff 001B</td>
<td>170</td>
<td>yes</td>
</tr>
<tr>
<td>Person67</td>
<td>Cardiff 001C</td>
<td>6</td>
<td>yes</td>
</tr>
<tr>
<td>Person67</td>
<td>Cardiff 009E</td>
<td>358</td>
<td>yes</td>
</tr>
<tr>
<td>Person79</td>
<td>Cardiff 002D</td>
<td>6</td>
<td>yes</td>
</tr>
<tr>
<td>Person80</td>
<td>Cardiff 003B</td>
<td>73</td>
<td>yes</td>
</tr>
<tr>
<td>Person81</td>
<td>Cardiff 004A</td>
<td>7</td>
<td>yes</td>
</tr>
<tr>
<td>Person82</td>
<td>Cardiff 004B</td>
<td>205</td>
<td>yes</td>
</tr>
<tr>
<td>Person83</td>
<td>Cardiff 005B</td>
<td>167</td>
<td>yes</td>
</tr>
<tr>
<td>Person84</td>
<td>Cardiff 006B</td>
<td>584</td>
<td>yes</td>
</tr>
<tr>
<td>Person85</td>
<td>Cardiff 006C</td>
<td>149</td>
<td>yes</td>
</tr>
<tr>
<td>Person96</td>
<td>Cardiff 007A</td>
<td>25</td>
<td>no</td>
</tr>
<tr>
<td>Person97</td>
<td>Cardiff 007D</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>Person98</td>
<td>Cardiff 007E</td>
<td>1</td>
<td>yes</td>
</tr>
<tr>
<td>Person09</td>
<td>Cardiff 008B</td>
<td>796</td>
<td>yes</td>
</tr>
<tr>
<td>Person09</td>
<td>Cardiff 008B</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>Person09</td>
<td>Cardiff 009A</td>
<td>202</td>
<td>yes</td>
</tr>
<tr>
<td>Person02</td>
<td>Cardiff 009A</td>
<td>21</td>
<td>yes</td>
</tr>
<tr>
<td>Person03</td>
<td>Cardiff 009B</td>
<td>843</td>
<td>yes</td>
</tr>
<tr>
<td>Person04</td>
<td>Cardiff 009C</td>
<td>704</td>
<td>yes</td>
</tr>
<tr>
<td>Person05</td>
<td>Cardiff 009C</td>
<td>82</td>
<td>yes</td>
</tr>
<tr>
<td>Person06</td>
<td>Cardiff 010A</td>
<td>214</td>
<td>yes</td>
</tr>
<tr>
<td>Person07</td>
<td>Cardiff 010B</td>
<td>715</td>
<td>yes</td>
</tr>
<tr>
<td>Person09</td>
<td>Cardiff 012C</td>
<td>31</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 012D</td>
<td>17</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 013D</td>
<td>13</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 014A</td>
<td>142</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 014C</td>
<td>76</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 014C</td>
<td>368</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 015A</td>
<td>589</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 015C</td>
<td>54</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 016D</td>
<td>521</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 017B</td>
<td>70</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 018A</td>
<td>781</td>
<td>yes</td>
</tr>
<tr>
<td>Person10</td>
<td>Cardiff 018B</td>
<td>19</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 018D</td>
<td>798</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 019D</td>
<td>68</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 019E</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 020A</td>
<td>272</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 020C</td>
<td>1058</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 020C</td>
<td>171</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 021A</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 021A</td>
<td>1826</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 021A</td>
<td>86</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 022A</td>
<td>38</td>
<td>yes</td>
</tr>
<tr>
<td>Person11</td>
<td>Cardiff 022C</td>
<td>772</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 022C</td>
<td>69</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 023A</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 001A</td>
<td>115</td>
<td>yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Cardiff 001B</td>
<td>17</td>
<td>yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Cardiff 001C</td>
<td>606</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 002A</td>
<td>31</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 002C</td>
<td>560</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 002C</td>
<td>11</td>
<td>yes</td>
</tr>
<tr>
<td>Person12</td>
<td>Cardiff 002C</td>
<td>210</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 002D</td>
<td>202</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 003A</td>
<td>1021</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 003B</td>
<td>330</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 003B</td>
<td>19</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 003C</td>
<td>112</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 003D</td>
<td>556</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 003D</td>
<td>28</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 004D</td>
<td>20</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 005C</td>
<td>5</td>
<td>yes</td>
</tr>
<tr>
<td>Person13</td>
<td>Cardiff 005D</td>
<td>28</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 005F</td>
<td>464</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 005G</td>
<td>337</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 005G</td>
<td>258</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 005G</td>
<td>68</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 005G</td>
<td>48</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 006D</td>
<td>6</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 007A</td>
<td>781</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 007C</td>
<td>85</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 008B</td>
<td>298</td>
<td>yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Cardiff 008C</td>
<td>5</td>
<td>yes</td>
</tr>
<tr>
<td>Person14</td>
<td>Cardiff 008D</td>
<td>479</td>
<td>yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Cardiff 009B</td>
<td>569</td>
<td>yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Cardiff 009B</td>
<td>10</td>
<td>yes</td>
</tr>
<tr>
<td>Person15</td>
<td>Cardiff 009B</td>
<td>1344</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009B</td>
<td>7</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009B</td>
<td>833</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>49</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>1035</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>31</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>7</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 009D</td>
<td>73</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 010C</td>
<td>151</td>
<td>yes</td>
</tr>
<tr>
<td>Person16</td>
<td>Cardiff 010C</td>
<td>263</td>
<td>yes</td>
</tr>
<tr>
<td>Person</td>
<td>Rhondda Cynon Taf 026A</td>
<td>Rhondda Cynon Taf 026A</td>
<td>818</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Person790</td>
<td>Rhondda Cynon Taf 025C</td>
<td>Rhondda Cynon Taf 027B</td>
<td>4</td>
</tr>
<tr>
<td>Person791</td>
<td>Rhondda Cynon Taf 027E</td>
<td>Rhondda Cynon Taf 027E</td>
<td>3</td>
</tr>
<tr>
<td>Person792</td>
<td>Rhondda Cynon Taf 027E</td>
<td>Rhondda Cynon Taf 027E</td>
<td>498</td>
</tr>
<tr>
<td>Person800</td>
<td>Rhondda Cynon Taf 027E</td>
<td>Rhondda Cynon Taf 027E</td>
<td>211</td>
</tr>
<tr>
<td>Person801</td>
<td>Rhondda Cynon Taf 030E</td>
<td>Rhondda Cynon Taf 030E</td>
<td>779</td>
</tr>
<tr>
<td>Person802</td>
<td>Rhondda Cynon Taf 027F</td>
<td>Rhondda Cynon Taf 027F</td>
<td>10</td>
</tr>
<tr>
<td>Person803</td>
<td>Rhondda Cynon Taf 027G</td>
<td>Rhondda Cynon Taf 027G</td>
<td>788</td>
</tr>
<tr>
<td>Person804</td>
<td>Rhondda Cynon Taf 028A</td>
<td>Rhondda Cynon Taf 028A</td>
<td>434</td>
</tr>
<tr>
<td>Person805</td>
<td>Rhondda Cynon Taf 028C</td>
<td>Cardiff 029A</td>
<td>1</td>
</tr>
<tr>
<td>Person806</td>
<td>Rhondda Cynon Taf 028E</td>
<td>Rhondda Cynon Taf 028E</td>
<td>13</td>
</tr>
<tr>
<td>Person807</td>
<td>Rhondda Cynon Taf 029C</td>
<td>Rhondda Cynon Taf 029C</td>
<td>195</td>
</tr>
<tr>
<td>Person808</td>
<td>Rhondda Cynon Taf 030A</td>
<td>Rhondda Cynon Taf 030A</td>
<td>6</td>
</tr>
<tr>
<td>Person809</td>
<td>Rhondda Cynon Taf 030A</td>
<td>Rhondda Cynon Taf 030A</td>
<td>519</td>
</tr>
<tr>
<td>Person810</td>
<td>Rhondda Cynon Taf 030C</td>
<td>Rhondda Cynon Taf 030C</td>
<td>229</td>
</tr>
<tr>
<td>Person811</td>
<td>Rhondda Cynon Taf 030D</td>
<td>Rhondda Cynon Taf 030D</td>
<td>146</td>
</tr>
<tr>
<td>Person812</td>
<td>Rhondda Cynon Taf 030E</td>
<td>Rhondda Cynon Taf 030E</td>
<td>99</td>
</tr>
<tr>
<td>Person813</td>
<td>Rhondda Cynon Taf 031A</td>
<td>Rhondda Cynon Taf 031A</td>
<td>1366</td>
</tr>
<tr>
<td>Person814</td>
<td>Rhondda Cynon Taf 031F</td>
<td>Rhondda Cynon Taf 031F</td>
<td>5</td>
</tr>
<tr>
<td>Person815</td>
<td>Swansea 001B</td>
<td>Swansea 001B</td>
<td>36</td>
</tr>
<tr>
<td>Person816</td>
<td>Swansea 001B</td>
<td>Swansea 001B</td>
<td>764</td>
</tr>
<tr>
<td>Person817</td>
<td>Swansea 001B</td>
<td>Swansea 001B</td>
<td>8</td>
</tr>
<tr>
<td>Person818</td>
<td>Swansea 001C</td>
<td>Swansea 001C</td>
<td>46</td>
</tr>
<tr>
<td>Person819</td>
<td>Swansea 002B</td>
<td>Swansea 002B</td>
<td>77</td>
</tr>
<tr>
<td>Person820</td>
<td>Swansea 002C</td>
<td>Swansea 002C</td>
<td>355</td>
</tr>
<tr>
<td>Person821</td>
<td>Swansea 002E</td>
<td>Swansea 002E</td>
<td>235</td>
</tr>
<tr>
<td>Person822</td>
<td>Swansea 003B</td>
<td>Swansea 003B</td>
<td>48</td>
</tr>
<tr>
<td>Person823</td>
<td>Swansea 003C</td>
<td>Swansea 003C</td>
<td>5</td>
</tr>
<tr>
<td>Person824</td>
<td>Swansea 003FE</td>
<td>Swansea 003FE</td>
<td>973</td>
</tr>
<tr>
<td>Person825</td>
<td>Swansea 003F</td>
<td>Swansea 003F</td>
<td>164</td>
</tr>
<tr>
<td>Person826</td>
<td>Swansea 004A</td>
<td>Swansea 004A</td>
<td>23</td>
</tr>
<tr>
<td>Person827</td>
<td>Swansea 004A</td>
<td>Swansea 004A</td>
<td>28</td>
</tr>
<tr>
<td>Person828</td>
<td>Swansea 004C</td>
<td>Swansea 004C</td>
<td>164</td>
</tr>
<tr>
<td>Person829</td>
<td>Swansea 004C</td>
<td>Swansea 004C</td>
<td>3</td>
</tr>
<tr>
<td>Person830</td>
<td>Swansea 004D</td>
<td>Swansea 004D</td>
<td>7</td>
</tr>
<tr>
<td>Person831</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>6</td>
</tr>
<tr>
<td>Person832</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>137</td>
</tr>
<tr>
<td>Person833</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>502</td>
</tr>
<tr>
<td>Person834</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>960</td>
</tr>
<tr>
<td>Person835</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>6</td>
</tr>
<tr>
<td>Person836</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>33</td>
</tr>
<tr>
<td>Person837</td>
<td>Swansea 004E</td>
<td>Swansea 004E</td>
<td>7</td>
</tr>
<tr>
<td>Person838</td>
<td>Swansea 005B</td>
<td>Swansea 005B</td>
<td>8</td>
</tr>
<tr>
<td>Person839</td>
<td>Swansea 005C</td>
<td>Swansea 005C</td>
<td>375</td>
</tr>
<tr>
<td>Person840</td>
<td>Swansea 005F</td>
<td>Swansea 005F</td>
<td>446</td>
</tr>
<tr>
<td>Person841</td>
<td>Swansea 006A</td>
<td>Swansea 006A</td>
<td>523</td>
</tr>
<tr>
<td>Person842</td>
<td>Swansea 006A</td>
<td>Swansea 006A</td>
<td>597</td>
</tr>
<tr>
<td>Person843</td>
<td>Swansea 006A</td>
<td>Swansea 006A</td>
<td>7</td>
</tr>
<tr>
<td>Person844</td>
<td>Swansea 007A</td>
<td>Swansea 007A</td>
<td>41</td>
</tr>
<tr>
<td>Person845</td>
<td>Swansea 007A</td>
<td>Swansea 007A</td>
<td>7</td>
</tr>
<tr>
<td>Person846</td>
<td>Swansea 007B</td>
<td>Swansea 007B</td>
<td>5</td>
</tr>
<tr>
<td>Person847</td>
<td>Swansea 007C</td>
<td>Swansea 007C</td>
<td>1170</td>
</tr>
<tr>
<td>Person848</td>
<td>Swansea 007C</td>
<td>Swansea 007C</td>
<td>255</td>
</tr>
<tr>
<td>Person849</td>
<td>Swansea 007E</td>
<td>Swansea 007E</td>
<td>6</td>
</tr>
<tr>
<td>Person850</td>
<td>Swansea 007E</td>
<td>Swansea 007E</td>
<td>887</td>
</tr>
<tr>
<td>Person851</td>
<td>Swansea 007E</td>
<td>Swansea 007E</td>
<td>661</td>
</tr>
<tr>
<td>Person852</td>
<td>Swansea 007E</td>
<td>Swansea 007E</td>
<td>21</td>
</tr>
<tr>
<td>Person853</td>
<td>Swansea 008A</td>
<td>Swansea 008A</td>
<td>7</td>
</tr>
<tr>
<td>Person854</td>
<td>Swansea 008A</td>
<td>Swansea 008A</td>
<td>119</td>
</tr>
<tr>
<td>Person855</td>
<td>Swansea 008B</td>
<td>Swansea 008B</td>
<td>806</td>
</tr>
<tr>
<td>Person856</td>
<td>Swansea 008B</td>
<td>Swansea 008B</td>
<td>406</td>
</tr>
<tr>
<td>Person857</td>
<td>Swansea 008B</td>
<td>Swansea 008B</td>
<td>1</td>
</tr>
<tr>
<td>Person858</td>
<td>Swansea 008C</td>
<td>Swansea 008C</td>
<td>23</td>
</tr>
<tr>
<td>Person859</td>
<td>Swansea 008C</td>
<td>Swansea 008C</td>
<td>212</td>
</tr>
<tr>
<td>Person860</td>
<td>Swansea 008D</td>
<td>Swansea 008D</td>
<td>4</td>
</tr>
<tr>
<td>Person861</td>
<td>Swansea 008D</td>
<td>Swansea 008D</td>
<td>363</td>
</tr>
<tr>
<td>Person862</td>
<td>Swansea 008D</td>
<td>Swansea 008D</td>
<td>18</td>
</tr>
<tr>
<td>Person863</td>
<td>Swansea 008E</td>
<td>Swansea 008E</td>
<td>9</td>
</tr>
<tr>
<td>Person864</td>
<td>Swansea 008E</td>
<td>Swansea 008E</td>
<td>111</td>
</tr>
<tr>
<td>Person865</td>
<td>Swansea 009B</td>
<td>Swansea 009B</td>
<td>5</td>
</tr>
<tr>
<td>Person866</td>
<td>Swansea 009B</td>
<td>Swansea 009B</td>
<td>9</td>
</tr>
<tr>
<td>Person867</td>
<td>Swansea 009C</td>
<td>Swansea 009C</td>
<td>25</td>
</tr>
<tr>
<td>Person868</td>
<td>Swansea 009C</td>
<td>Swansea 009C</td>
<td>34</td>
</tr>
<tr>
<td>Person869</td>
<td>Swansea 009C</td>
<td>Swansea 009C</td>
<td>462</td>
</tr>
<tr>
<td>Person870</td>
<td>Swansea 009D</td>
<td>Swansea 009D</td>
<td>238</td>
</tr>
<tr>
<td>Person871</td>
<td>Swansea 009D</td>
<td>Swansea 009D</td>
<td>12</td>
</tr>
<tr>
<td>Person872</td>
<td>Swansea 009E</td>
<td>Swansea 009E</td>
<td>8</td>
</tr>
<tr>
<td>Person873</td>
<td>Swansea 009F</td>
<td>Swansea 009F</td>
<td>38</td>
</tr>
<tr>
<td>Person874</td>
<td>Swansea 009F</td>
<td>Swansea 009F</td>
<td>73</td>
</tr>
<tr>
<td>Person875</td>
<td>Swansea 010A</td>
<td>Swansea 010A</td>
<td>249</td>
</tr>
<tr>
<td>Person876</td>
<td>Swansea 010A</td>
<td>Swansea 010A</td>
<td>266</td>
</tr>
<tr>
<td>Person877</td>
<td>Swansea 010A</td>
<td>Swansea 010A</td>
<td>5</td>
</tr>
<tr>
<td>Person878</td>
<td>Swansea 010B</td>
<td>Swansea 010B</td>
<td>12</td>
</tr>
<tr>
<td>Person879</td>
<td>Swansea 010C</td>
<td>Swansea 010C</td>
<td>170</td>
</tr>
<tr>
<td>Person880</td>
<td>Swansea 010D</td>
<td>Swansea 010D</td>
<td>32</td>
</tr>
<tr>
<td>Person881</td>
<td>Swansea 010D</td>
<td>Swansea 010D</td>
<td>104</td>
</tr>
<tr>
<td>Person882</td>
<td>Swansea 010D</td>
<td>Swansea 010D</td>
<td>1242</td>
</tr>
<tr>
<td>Person883</td>
<td>Swansea 010D</td>
<td>Swansea 010D</td>
<td>3</td>
</tr>
<tr>
<td>Person884</td>
<td>Swansea 010D</td>
<td>Swansea 010D</td>
<td>14</td>
</tr>
<tr>
<td>Person885</td>
<td>Swansea 010D</td>
<td>Swansea 010D</td>
<td>15</td>
</tr>
</tbody>
</table>
Appendix D – Health risk co-behaviour codes

Tables D.1 to D.4 illustrate the codes with a division of gender that were present in the health co-behavioural tweets before they were cut down to three core theme per co-behaviour reference.

Table D.1 Content analysis of tweets with both a smoking and alcohol use reference

<table>
<thead>
<tr>
<th>Code</th>
<th>Categories</th>
<th>Boys (n=400)</th>
<th>Girls (n=369)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alcohol use (n=769)</td>
<td>194</td>
<td>140</td>
<td>“living the dream at the moment, sun’s out, BBQ’s out, beers out, cigarette's out, flip flops out #brill!”</td>
</tr>
<tr>
<td></td>
<td>Having a smoke and alcohol together (n=334)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Wanting both alcohol and smoking (n=55)</td>
<td>21</td>
<td>34</td>
<td>“I need a double gin/ triple gin/ a bottle of gin and a cigarette. ASAP.”</td>
</tr>
<tr>
<td>3</td>
<td>Buying them (n=34)</td>
<td>13</td>
<td>21</td>
<td>“Beer and fags. Shopping list sorted.”</td>
</tr>
<tr>
<td>4</td>
<td>Smoking when drinking (n=38)</td>
<td>19</td>
<td>19</td>
<td>“always need a fag with a drink”</td>
</tr>
<tr>
<td>5</td>
<td>Quit smoking just not when drinking alcohol (n=28)</td>
<td>15</td>
<td>13</td>
<td>“no way! I quit smoking except for when I'm drinking”</td>
</tr>
<tr>
<td>6</td>
<td>Quit smoking but not alcohol (n=23)</td>
<td>10</td>
<td>13</td>
<td>“Ever since I stopped smoking I've became an alcoholic”</td>
</tr>
<tr>
<td>7</td>
<td>Quitting smoking and quitting alcohol (n=85)</td>
<td>39</td>
<td>46</td>
<td>“GS has it right, stopping drinking and smoking won't make me live longer, it'll just seem longer”</td>
</tr>
<tr>
<td>8</td>
<td>Remembering an occasion (n=21)</td>
<td>11</td>
<td>10</td>
<td>“Still can’t get over this picture. Me with a fag and a vodka and coke at the age of 4, nothing has changed??”</td>
</tr>
<tr>
<td>9</td>
<td>Comparing smoking and alcohol (n=54)</td>
<td>28</td>
<td>26</td>
<td>“The police should focus more on underage drinking than people smoking weed, I'm just sitting here chilling man?”</td>
</tr>
<tr>
<td>10</td>
<td>Facts about alcohol and smoking (n=22)</td>
<td>13</td>
<td>9</td>
<td>“The health risks of being obese are worse than the risks of heavy drinking, smoking and poverty.”</td>
</tr>
<tr>
<td>11</td>
<td>Other (n=74)</td>
<td>37</td>
<td>37</td>
<td>“My grandma just told me to make sure people don’t put weed or something in my drink when I'm on holiday ???? bless her”</td>
</tr>
</tbody>
</table>
Table D.2 Content analysis of tweets with both a smoking and dieting & healthy eating reference

<table>
<thead>
<tr>
<th>Code</th>
<th>Categories</th>
<th>Boys (n=212)</th>
<th>Girls (n=259)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dieting &amp; healthy eating (n=371)</td>
<td>Smoking &amp; eating (n=183)</td>
<td>99</td>
<td>84</td>
<td>Sex, cigarette, food, Xbox, sex and sex.”</td>
</tr>
<tr>
<td>2</td>
<td>Healthy lifestyle (n=89)</td>
<td>21</td>
<td>68</td>
<td>’New years’ resolutions - reach 7stone, quit smoking for good, study more, drink less fizzy drinks, save money...’</td>
</tr>
<tr>
<td>3</td>
<td>Quit smoking but want to eat unhealthy (n=79)</td>
<td>35</td>
<td>44</td>
<td>“Day 3 of no cigs - I’m eating everything in sight?”</td>
</tr>
<tr>
<td>4</td>
<td>Healthy eating but smoking (n=13)</td>
<td>5</td>
<td>8</td>
<td>”Got to get out the habit of doing exercise, then going for a fag and then eating heavily buttered toast, defeats the purpose really.”</td>
</tr>
<tr>
<td>5</td>
<td>Losing weight through smoking (n=10)</td>
<td>4</td>
<td>6</td>
<td>”I lost weight from taking up smoking and vodka enemas”</td>
</tr>
<tr>
<td>6</td>
<td>Food flavoured smoking/ smoking flavoured food (n=9)</td>
<td>6</td>
<td>3</td>
<td>”Also white chocolate &amp; raspberry iced coffees taste like hookah”</td>
</tr>
<tr>
<td>7</td>
<td>Comparing unhealthy eating and smoking (n=20)</td>
<td>10</td>
<td>10</td>
<td>”There should be severe health warnings on fast food and alcohol as there is on cigarettes ?? no ??”</td>
</tr>
<tr>
<td>8</td>
<td>Buying one instead of the other (n=23)</td>
<td>13</td>
<td>10</td>
<td>”Go to the shop for tobacco and I come back with crisps, sweets and chocolate #suchachild”</td>
</tr>
<tr>
<td>9</td>
<td>Facts about smoking (n=14)</td>
<td>9</td>
<td>5</td>
<td>”People who smoke marijuana are less likely to be obese!”</td>
</tr>
<tr>
<td>10</td>
<td>People smoking while other eat (n=9)</td>
<td>1</td>
<td>8</td>
<td>”Don't care what anyone says, smoking around people that are eating is just rude???”</td>
</tr>
<tr>
<td>11</td>
<td>Other (n=20)</td>
<td>9</td>
<td>11</td>
<td>”Wish losing weight was as easy as ash dropping off a fag”</td>
</tr>
</tbody>
</table>
### Table D.3 Content analysis of tweets with both a smoking and physical exercise reference

<table>
<thead>
<tr>
<th>Code</th>
<th>Categories</th>
<th>Boys (n=86)</th>
<th>Girls (n=108)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Physical Exercise (n= 194)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Smoking &amp; exercising (n=26)</td>
<td>14</td>
<td>12</td>
<td>‘My worst habit is smoking a cigarette straight after the gym, defeats the object’</td>
</tr>
<tr>
<td>3</td>
<td>Healthy lifestyle (n=108)</td>
<td>47</td>
<td>61</td>
<td>‘2 and half weeks without smoking! Doing a charity run today as well! #LifesGood #QuitSmoking #CharityRun’</td>
</tr>
<tr>
<td>4</td>
<td>Bad fitness through smoking (n=24)</td>
<td>9</td>
<td>15</td>
<td>‘Was DYING earlier in gym... Defo stopping smoking jheeezzz’</td>
</tr>
<tr>
<td>5</td>
<td>Smoking but no exercise (n=10)</td>
<td>5</td>
<td>5</td>
<td>‘Everyone’s on a fitness hype, I ate a microwave pizza for breakfast then a curry for dinner and im about to have a fag #fuckofffitness’</td>
</tr>
<tr>
<td>6</td>
<td>Comparing smoking to exercise (n=4)</td>
<td>3</td>
<td>1</td>
<td>‘I’d much rather go to the gym than have a smoking habit’</td>
</tr>
<tr>
<td>7</td>
<td>Others smoking and exercising (n=16)</td>
<td>6</td>
<td>10</td>
<td>‘People that stand outside the gym smoking.. Are you for real?!’</td>
</tr>
<tr>
<td>8</td>
<td>Other (n=6)</td>
<td>2</td>
<td>4</td>
<td>‘i can smell smoke in my hair, just from running past someone smoking....vile vile’</td>
</tr>
</tbody>
</table>

### Table D.4 Content analysis of tweets with both a smoking and combined health behaviours (i.e. alcohol use, dieting & healthy eating, and exercise & fitness)

<table>
<thead>
<tr>
<th>Code</th>
<th>Categories</th>
<th>Boys (n=48)</th>
<th>Girls (n=57)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Getting healthy by doing it all (n=42)</td>
<td>19</td>
<td>23</td>
<td>‘2 weeks of healthy eating, 1 day of no smoking and a cheeky run to top things off. Can’t wait for gym tomo!! :) #focused’</td>
</tr>
<tr>
<td>2</td>
<td>Unhealthy combinations (n=27)</td>
<td>12</td>
<td>15</td>
<td>‘My body hates me from too much drinking, smoking and eating #holidaycomedown’</td>
</tr>
<tr>
<td>3</td>
<td>Doing one thing but not the other (n=12)</td>
<td>4</td>
<td>8</td>
<td>‘It’s all good going to the gym and eating moderately healthily but I ruin it all by smoking and drinking countless bottles of wine ??’</td>
</tr>
<tr>
<td>4</td>
<td>New Year’s resolutions/ Stoptober /Lent (n=8)</td>
<td>1</td>
<td>7</td>
<td>‘New years resolutions - reach 7stone, quit smoking for good, study more, drink less fizzy drinks, save money...’</td>
</tr>
<tr>
<td>5</td>
<td>Comparing one to the other (n=8)</td>
<td>7</td>
<td>1</td>
<td>‘there should be severe health warnings on fast food and alcohol as there is on cigarettes ?? no ??’</td>
</tr>
<tr>
<td>6</td>
<td>Financial comments on more health behaviours (n=5)</td>
<td>3</td>
<td>2</td>
<td>‘I’ve spent £60 since Christmas to now , mainly on food, alcohol and fags #help’</td>
</tr>
<tr>
<td>7</td>
<td>Other (n=3)</td>
<td>2</td>
<td>1</td>
<td>‘Rolled a fag but didn’t have time to smoke it before my food came. My hot plate was on it so it was dead. So I threw it. It went in my wine.’</td>
</tr>
</tbody>
</table>
### Appendix E – Table of the description of 50 subsample members

This table includes the altered names and profile information of the individuals that were randomly selected for the in-depth examination of the smoking tweets in Chapter 9.

Table F.1 The details of the 50 subsample members selected for the analyses in Chapter 9.

<table>
<thead>
<tr>
<th>Alternative name</th>
<th>Gender</th>
<th>Tweets</th>
<th>Following</th>
<th>Followers</th>
<th>Likes</th>
<th>Joined</th>
<th>Age</th>
<th>Smoking status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaron</td>
<td>1</td>
<td>12800</td>
<td>408</td>
<td>395</td>
<td>116</td>
<td>12-jan-10</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Alexander</td>
<td>1</td>
<td>3750</td>
<td>719</td>
<td>1027</td>
<td>2657</td>
<td>10-jan-10</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Benedict</td>
<td>1</td>
<td>2902</td>
<td>612</td>
<td>284</td>
<td>30</td>
<td>3-jan-08</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Bernard</td>
<td>1</td>
<td>2028</td>
<td>263</td>
<td>255</td>
<td>468</td>
<td>9-jan-12</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Catherine</td>
<td>2</td>
<td>5586</td>
<td>67</td>
<td>465</td>
<td>855</td>
<td>12-jan-12</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Chanel</td>
<td>2</td>
<td>1035</td>
<td>269</td>
<td>106</td>
<td>607</td>
<td>7-jan-12</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Diana</td>
<td>2</td>
<td>8741</td>
<td>165</td>
<td>295</td>
<td>458</td>
<td>7-jan-09</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Dora</td>
<td>2</td>
<td>475</td>
<td>1967</td>
<td>722</td>
<td>26</td>
<td>10-jan-09</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Edmund</td>
<td>1</td>
<td>5582</td>
<td>391</td>
<td>484</td>
<td>1517</td>
<td>9-jan-11</td>
<td>20 smoker</td>
<td></td>
</tr>
<tr>
<td>Emilia</td>
<td>2</td>
<td>4145</td>
<td>360</td>
<td>221</td>
<td>672</td>
<td>7-jan-12</td>
<td>relapser</td>
<td></td>
</tr>
<tr>
<td>Francis</td>
<td>1</td>
<td>4997</td>
<td>1202</td>
<td>1245</td>
<td>1478</td>
<td>5-jan-12</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Freddy</td>
<td>1</td>
<td>2829</td>
<td>380</td>
<td>264</td>
<td>19</td>
<td>4-jan-10</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Gary</td>
<td>1</td>
<td>13700</td>
<td>1262</td>
<td>484</td>
<td>512</td>
<td>6-jan-09</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>2</td>
<td>3106</td>
<td>416</td>
<td>469</td>
<td>424</td>
<td>10-jan-11</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Hayley</td>
<td>2</td>
<td>11200</td>
<td>501</td>
<td>453</td>
<td>2666</td>
<td>6-jan-12</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Holly</td>
<td>2</td>
<td>12800</td>
<td>513</td>
<td>696</td>
<td>5275</td>
<td>1-jan-11</td>
<td>20 non-smoker</td>
<td></td>
</tr>
<tr>
<td>Isaac</td>
<td>1</td>
<td>1794</td>
<td>275</td>
<td>299</td>
<td>203</td>
<td>6-jan-12</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Isabella</td>
<td>2</td>
<td>850</td>
<td>283</td>
<td>216</td>
<td>714</td>
<td>11-jan-12</td>
<td>24 smoker</td>
<td></td>
</tr>
<tr>
<td>Jessica</td>
<td>2</td>
<td>15500</td>
<td>338</td>
<td>505</td>
<td>6931</td>
<td>9-jan-09</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Juliette</td>
<td>2</td>
<td>591</td>
<td>409</td>
<td>202</td>
<td>73</td>
<td>8-jan-10</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Kevin</td>
<td>1</td>
<td>10359</td>
<td>348</td>
<td>337</td>
<td>3257</td>
<td>2-jan-10</td>
<td>relapser</td>
<td></td>
</tr>
<tr>
<td>Kyle</td>
<td>1</td>
<td>1841</td>
<td>635</td>
<td>589</td>
<td>773</td>
<td>5-jan-12</td>
<td>20 non-smoker</td>
<td></td>
</tr>
<tr>
<td>Lauren</td>
<td>2</td>
<td>5360</td>
<td>688</td>
<td>109</td>
<td>706</td>
<td>7-jan-09</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Lucas</td>
<td>1</td>
<td>9494</td>
<td>217</td>
<td>347</td>
<td>8720</td>
<td>11-jan-11</td>
<td>18 non-smoker</td>
<td></td>
</tr>
<tr>
<td>Maxwell</td>
<td>1</td>
<td>620</td>
<td>105</td>
<td>87</td>
<td>90</td>
<td>11-jan-09</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Megan</td>
<td>2</td>
<td>12500</td>
<td>464</td>
<td>536</td>
<td>4317</td>
<td>6-jan-10</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Nicole</td>
<td>2</td>
<td>222</td>
<td>51</td>
<td>30</td>
<td>81</td>
<td>3-jan-10</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Nina</td>
<td>2</td>
<td>17900</td>
<td>464</td>
<td>439</td>
<td>2339</td>
<td>6-jan-09</td>
<td>23 smoker</td>
<td></td>
</tr>
<tr>
<td>Olaf</td>
<td>1</td>
<td>3976</td>
<td>507</td>
<td>635</td>
<td>1087</td>
<td>6-jan-10</td>
<td>25 quitter</td>
<td></td>
</tr>
<tr>
<td>Oliver</td>
<td>1</td>
<td>1486</td>
<td>632</td>
<td>314</td>
<td>1133</td>
<td>3-jan-12</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Paige</td>
<td>2</td>
<td>10500</td>
<td>83</td>
<td>225</td>
<td>407</td>
<td>12-jan-10</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Pippa</td>
<td>2</td>
<td>12400</td>
<td>382</td>
<td>483</td>
<td>2</td>
<td>10-jan-10</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Quentin</td>
<td>1</td>
<td>7142</td>
<td>1577</td>
<td>1974</td>
<td>1670</td>
<td>7-jan-09</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Quint</td>
<td>1</td>
<td>8942</td>
<td>535</td>
<td>523</td>
<td>3852</td>
<td></td>
<td></td>
<td>non-smoker</td>
</tr>
<tr>
<td>Rachel</td>
<td>2</td>
<td>17800</td>
<td>338</td>
<td>453</td>
<td>4776</td>
<td>7-jan-10</td>
<td>23 smoker</td>
<td></td>
</tr>
<tr>
<td>Rebecca</td>
<td>2</td>
<td>13600</td>
<td>340</td>
<td>1375</td>
<td>6987</td>
<td>4-jan-12</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Sarah</td>
<td>2</td>
<td>6419</td>
<td>342</td>
<td>158</td>
<td>666</td>
<td>11-jan-11</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Stephen</td>
<td>1</td>
<td>1454</td>
<td>104</td>
<td>134</td>
<td>263</td>
<td>7-jan-12</td>
<td>20 non-smoker</td>
<td></td>
</tr>
<tr>
<td>Taye</td>
<td>1</td>
<td>33100</td>
<td>474</td>
<td>423</td>
<td>33</td>
<td>4-jan-10</td>
<td>21 smoker</td>
<td></td>
</tr>
<tr>
<td>Tess</td>
<td>2</td>
<td>334</td>
<td>131</td>
<td>66</td>
<td>301</td>
<td>3-jan-13</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Uma</td>
<td>2</td>
<td>14800</td>
<td>894</td>
<td>500</td>
<td>1498</td>
<td>4-jan-09</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Usain</td>
<td>1</td>
<td>1399</td>
<td>285</td>
<td>195</td>
<td>425</td>
<td>9-jan-12</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Victoria</td>
<td>2</td>
<td>5762</td>
<td>386</td>
<td>489</td>
<td>2365</td>
<td>6-jan-12</td>
<td>smoker</td>
<td></td>
</tr>
<tr>
<td>Vivian</td>
<td>2</td>
<td>609</td>
<td>78</td>
<td>18</td>
<td>90</td>
<td>3-jan-09</td>
<td>quitter</td>
<td></td>
</tr>
<tr>
<td>Wendy</td>
<td>2</td>
<td>23695</td>
<td>967</td>
<td>1032</td>
<td>9501</td>
<td>10-jan-11</td>
<td>non-smoker</td>
<td></td>
</tr>
<tr>
<td>Whitney</td>
<td>2</td>
<td>814</td>
<td>819</td>
<td>140</td>
<td>182</td>
<td>9-jan-10</td>
<td>29 quitter</td>
<td></td>
</tr>
<tr>
<td>Yara</td>
<td>2</td>
<td>1527</td>
<td>213</td>
<td>178</td>
<td>85</td>
<td>7-jan-12</td>
<td>27 quitter</td>
<td></td>
</tr>
<tr>
<td>Yvonne</td>
<td>2</td>
<td>5002</td>
<td>191</td>
<td>263</td>
<td>1087</td>
<td>7-jan-14</td>
<td>relapser</td>
<td></td>
</tr>
<tr>
<td>Zachary</td>
<td>1</td>
<td>9210</td>
<td>215</td>
<td>516</td>
<td>1721</td>
<td>11-jan-09</td>
<td>23 smoker</td>
<td></td>
</tr>
<tr>
<td>Zander</td>
<td>1</td>
<td>8702</td>
<td>469</td>
<td>1178</td>
<td>1420</td>
<td>10-jan-11</td>
<td>22 relapser</td>
<td></td>
</tr>
</tbody>
</table>
Appendix F – Codes in Atlas.TI

Table E.1 Description of the inductive process of analysing and layering the Twitter archives of the randomly selected 50 sample members.

<table>
<thead>
<tr>
<th>Initial codes (at first read through)</th>
<th>Transformed codes to fit smoking context</th>
<th>Final product in the thesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversations</td>
<td>keep in touch with friends and family</td>
<td>Socialising</td>
</tr>
<tr>
<td>Congratulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sports commentary</td>
<td>communicate more easily and with more people at the same time</td>
<td></td>
</tr>
<tr>
<td>Irritations</td>
<td>express myself freely</td>
<td></td>
</tr>
<tr>
<td>Tirades</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boredom</td>
<td>Could be: be entertained, relax, see what others are up to, pass the time, get information and advice, and learn interesting things</td>
<td>General interest</td>
</tr>
<tr>
<td>Indeterminate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking tweets</td>
<td>Day of the smoking tweets</td>
<td>Possible motivation for smoking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Possible motivation for quitting</td>
</tr>
<tr>
<td></td>
<td>Tobacco</td>
<td>Show difference of tweeting about each type</td>
</tr>
<tr>
<td></td>
<td>Marijuana</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E-cigarettes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shisha</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Smokers</td>
<td>Show difference of tweeting within status</td>
</tr>
<tr>
<td></td>
<td>Quitters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relapsers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-smokers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stigmatizing content</td>
<td>Show stigmatisation of smokers and self-stigmatisation</td>
</tr>
<tr>
<td>Contact with The Filter Wales (TFW)</td>
<td>TFW tweets were added into the Twitter feeds</td>
<td>Show the TFW interaction</td>
</tr>
</tbody>
</table>
Appendix G – Twitter Style analysis tables

These tables are the output of the AnalyzeWords analysis of the entire Twitter history (tweets and profile) of the subsample members of Chapter 9.

Table G.1 The number of young people for the different scores in the Upbeat linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Upbeat</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>2 (12.5%)</td>
<td>3 (18.8%)</td>
<td>10 (62.5%)</td>
<td>1 (6.3%)</td>
<td></td>
<td>16 (100%)</td>
</tr>
<tr>
<td>Smokers</td>
<td></td>
<td>2 (22.2%)</td>
<td>5 (55.6%)</td>
<td>2 (22.2%)</td>
<td></td>
<td>9 (100%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>1 (8.3%)</td>
<td>1 (8.3%)</td>
<td>6 (50.0%)</td>
<td>2 (16.7%)</td>
<td>2 (16.7%)</td>
<td>12 (100%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>2 (66.7%)</td>
<td>1 (33.3%)</td>
<td></td>
<td></td>
<td></td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>3 (7.5%)</td>
<td>8 (20%)</td>
<td>22 (55%)</td>
<td>5 (12.5%)</td>
<td>2 (5%)</td>
<td>40 (100%)</td>
</tr>
</tbody>
</table>

Table G.2 The number of young people for the different scores in the Worried linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Worried</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>6 (37.5%)</td>
<td>7 (43.8%)</td>
<td>1 (6.3%)</td>
<td>2 (12.5%)</td>
<td></td>
<td>16 (100%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>1 (11.1%)</td>
<td>3 (33.3%)</td>
<td>3 (33.3%)</td>
<td>2 (22.2%)</td>
<td></td>
<td>9 (100%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>4 (33.3%)</td>
<td>3 (25%)</td>
<td>3 (25%)</td>
<td>2 (16.7%)</td>
<td></td>
<td>12 (100%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>2 (66.7%)</td>
<td>1 (33.3%)</td>
<td></td>
<td></td>
<td></td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>1 (2.5%)</td>
<td>15 (37.5%)</td>
<td>14 (35%)</td>
<td>4 (10%)</td>
<td>6 (15%)</td>
<td>40 (100%)</td>
</tr>
</tbody>
</table>

Table G.3 The number of young people for the different scores in the Angry linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Angry</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers Smokers</td>
<td>5 (31.3%)</td>
<td>6 (37.5%)</td>
<td>4 (25.0%)</td>
<td>1 (6.3%)</td>
<td></td>
<td>16 (100%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>1 (11.1%)</td>
<td>1 (11.1%)</td>
<td>2 (22.2%)</td>
<td>5 (55.6%)</td>
<td></td>
<td>9 (100%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>4 (33.3%)</td>
<td>4 (33.3%)</td>
<td>3 (25.0%)</td>
<td>1 (8.3%)</td>
<td></td>
<td>12 (100%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td></td>
<td></td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>11 (27.5%)</td>
<td>12 (30%)</td>
<td>10 (25%)</td>
<td>7 (17.5%)</td>
<td></td>
<td>40 (100%)</td>
</tr>
</tbody>
</table>
Table G.4 The number of young people for the different scores in the ‘Depressed’ linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>1 (6.3%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>2 (22.2%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>5 (41.7%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>6 (15%)</td>
</tr>
</tbody>
</table>

Table G.5 The number of young people for the different scores in the Plugged-in linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Plugged-in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>1 (6.3%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>2 (12.5%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>1 (11.1%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>2 (16.7%)</td>
</tr>
<tr>
<td>Total</td>
<td>1 (2.5%)</td>
</tr>
</tbody>
</table>

Table G.6 The number of young people for the different scores in the Personable linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Personable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>4 (25.0%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>2 (22.2%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>3 (25.0%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>1 (33.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>9 (22.5%)</td>
</tr>
</tbody>
</table>
### Table G.7 The number of young people for the different scores in the Arrogant/Distant linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Arrogant/Distant</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>2 (12.5%)</td>
<td>6 (37.5%)</td>
<td>7 (43.8%)</td>
<td>1 (6.3%)</td>
<td>16 (100%)</td>
<td></td>
</tr>
<tr>
<td>Smokers</td>
<td>1 (11.1%)</td>
<td>5 (55.6%)</td>
<td>3 (33.3%)</td>
<td></td>
<td>9 (100%)</td>
<td></td>
</tr>
<tr>
<td>Quitters</td>
<td>1 (8.3%)</td>
<td>10 (83.3%)</td>
<td>1 (8.3%)</td>
<td></td>
<td>12 (100%)</td>
<td></td>
</tr>
<tr>
<td>Relapsers</td>
<td>3 (100%)</td>
<td></td>
<td></td>
<td></td>
<td>3 (100%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3 (7.5%)</td>
<td>15 (37.5%)</td>
<td>20 (50%)</td>
<td>2 (5%)</td>
<td>40 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

### Table G.8 The number of young people for the different scores in the Spacey/Valley woman linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Spacy/ Valley Girl</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>9 (56.3%)</td>
<td>7 (43.8)</td>
<td></td>
<td></td>
<td>16 (100%)</td>
<td></td>
</tr>
<tr>
<td>Smokers</td>
<td>4 (44.4%)</td>
<td>5 (55.6%)</td>
<td></td>
<td></td>
<td>9 (100%)</td>
<td></td>
</tr>
<tr>
<td>Quitters</td>
<td>1 (8.3%)</td>
<td>8 (66.7%)</td>
<td>3 (25.0%)</td>
<td></td>
<td>12 (100%)</td>
<td></td>
</tr>
<tr>
<td>Relapsers</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td></td>
<td>3 (100%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2 (5%)</td>
<td>22 (55%)</td>
<td>16 (40%)</td>
<td></td>
<td>40 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

### Table G.9 The number of young people for the different scores in the Analytic linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Analytic</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>4 (25.0%)</td>
<td>4 (25.0%)</td>
<td>6 (37.5%)</td>
<td>2 (12.5%)</td>
<td>16 (100%)</td>
<td></td>
</tr>
<tr>
<td>Smokers</td>
<td>1 (11.1%)</td>
<td>1 (11.1%)</td>
<td>2 (22.2%)</td>
<td>3 (33.3%)</td>
<td>9 (100%)</td>
<td></td>
</tr>
<tr>
<td>Quitters</td>
<td>4 (33.3%)</td>
<td>3 (25.0%)</td>
<td>5 (41.7%)</td>
<td></td>
<td>12 (100%)</td>
<td></td>
</tr>
<tr>
<td>Relapsers</td>
<td>2 (66.7%)</td>
<td>1 (33.3%)</td>
<td></td>
<td></td>
<td>3 (100%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1 (2.5%)</td>
<td>9 (22.5%)</td>
<td>11 (27.5%)</td>
<td>15 (37.5%)</td>
<td>2 (10%)</td>
<td>40 (100%)</td>
</tr>
</tbody>
</table>
Table G.10 The number of young people for the different scores in the Sensory linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Sensory</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>2 (12.5%)</td>
<td>2 (12.5%)</td>
<td>9 (56.3%)</td>
<td>3 (18.8%)</td>
<td>16 (100%)</td>
<td></td>
</tr>
<tr>
<td>Smokers</td>
<td></td>
<td>2 (22.2%)</td>
<td>2 (22.2%)</td>
<td>5 (55.6%)</td>
<td>9 (100%)</td>
<td></td>
</tr>
<tr>
<td>Quitters</td>
<td>2 (16.7%)</td>
<td>1 (8.3%)</td>
<td>2 (16.7%)</td>
<td>5 (41.7%)</td>
<td>12 (100%)</td>
<td></td>
</tr>
<tr>
<td>Relapsers</td>
<td></td>
<td>3 (100%)</td>
<td></td>
<td></td>
<td>3 (100%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2 (5%)</td>
<td>3 (7.5%)</td>
<td>6 (15%)</td>
<td>13 (32.5%)</td>
<td>16 (40%)</td>
<td></td>
</tr>
</tbody>
</table>

Table G.11 The number of young people for the different scores in the In-the-moment linguistic analysis

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>In-the-moment</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very low (%)</td>
<td>Low (%)</td>
<td>Average (%)</td>
<td>High (%)</td>
<td>Very high (%)</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>1 (6.3%)</td>
<td>1 (6.3%)</td>
<td>5 (31.3%)</td>
<td>8 (50.0%)</td>
<td>1 (6.3%)</td>
<td>16 (100%)</td>
</tr>
<tr>
<td>Smokers</td>
<td></td>
<td>5 (55.6%)</td>
<td>4 (44.4%)</td>
<td></td>
<td></td>
<td>9 (100%)</td>
</tr>
<tr>
<td>Quitters</td>
<td>2 (16.7%)</td>
<td>6 (50.0%)</td>
<td>4 (33.3%)</td>
<td></td>
<td></td>
<td>12 (100%)</td>
</tr>
<tr>
<td>Relapsers</td>
<td>1 (33.3%)</td>
<td>2 (66.7%)</td>
<td></td>
<td></td>
<td></td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>1 (2.5%)</td>
<td>3 (7.5%)</td>
<td>17 (42.5%)</td>
<td>18 (45%)</td>
<td>1 (2.5%)</td>
<td>40 (100%)</td>
</tr>
</tbody>
</table>