

**Socio-economic disparities in science knowledge,  
biomedical self-efficacy, and public participation  
in medical decision-making**

Andreea-Loredana Moldovan

A thesis submitted for the degree of Doctor of Philosophy

Department of Sociology

University of Essex

February 2018

## **Declarations**

No part of this thesis has been submitted for another degree. All the work is original and it is my own (as described below).

Chapter 3 is co-authored with Professor Nick Allum and will soon be submitted to Science Communication.

The other two chapters are exclusively my own work. The second chapter is under review for publication in the International Journal of Social Research Methodology. The fourth chapter has been submitted to Social Science Research.



## **Acknowledgements**

I would like to thank my principal supervisor here at Essex, Professor Nick Allum, for his support and guidance all these years. I would also like to thank him for introducing me to the field of Public Understanding of Science, which has opened a lot of doors for me.

I would also like to thank Dr Hilary Leever, my supervisor at the Wellcome Trust, for her guidance and for treating me as an integral part of her Education & Learning Team during the periods I spent at the Wellcome Trust.

I have been fortunate to form great friendships throughout my PhD. A special thank you to Can Zuo and Eleana Nikiforidou, with whom I shared office 5A.309. It has been an absolute pleasure to know you both.

This PhD was supported by a 1+3 ESRC scholarship (grant number ES/J500045/1), in collaboration with the Wellcome Trust. I am very grateful for the opportunity I have been given. I have also benefited from the opportunity to spend a month at the Danish Centre for Studies in Research and Research Policy. I thank Dr Mejlgaard for inviting me and the ESRC for funding my visit.

I would like to thank my parents, Alina and Dan, and my sister, Raluca, for their endless love and support.

Finally, I want to thank my husband, Matt, for his seemingly endless patience, for proofreading a previous version of my PhD, and for the interesting and thought-provoking conversations regarding my research.



## Summary

The thesis consists of three self-contained articles that empirically investigate socio-economic differences in, and interrelationships amongst, science knowledge, biomedical self-efficacy, and participation in medical decision-making.

Chapter 2 investigates age-related bias in the science knowledge questions in the Wellcome Trust Monitor Survey Waves I and II. It also examines what evidence there is for three dimensions of knowledge. Chapter 3 studies the influence of Internet use and paying attention to medical stories online in reducing science knowledge and biomedical self-efficacy gaps between low and high educational groups. Wave II of the Wellcome Trust Monitor Surveys is employed in this chapter. Chapter 4 scrutinises the influence of various socio-economic factors, biomedical self-efficacy, and trust in physicians and other medical practitioners on public willingness and confidence to take part in the medical decision-making process. Chapter 4 uses Wave III of the Wellcome Trust Monitor Survey.

Chapter 2 finds evidence for age-related bias in the science knowledge battery of questions; no evidence of a misinformed group of respondents was found; a group who consistently said they didn't know instead of providing an answer that was wrong was found; a sensitivity analysis showed that using the summed score approach leads to the same substantive conclusions as a model taking into account age-related non-invariance. Chapter 3 finds evidence of education-based knowledge and efficacy gaps. It also finds some evidence that the Internet can help reduce that democratic deficit in information. Chapter 4 finds evidence that people are generally confident to participate. Those who are more self-efficacious are also more confident to participate in medical decisions. The opposite held true for those who place high trust in doctors. Women were found to be more confident than men.



# Table of Contents

<b>1. Overall Introduction.....</b>	<b>1</b>
<b>1.1 Thesis background .....</b>	<b>1</b>
<b>1.2 Breakdown of chapters.....</b>	<b>4</b>
1.2.1 Publication plan:.....	7
<b>1.3 Collaboration with the Wellcome Trust .....</b>	<b>7</b>
<b>2. Information, Misinformation, Lack of Information, and Age-based Differences in Science Knowledge: A Multi-group Latent Class Model. ....</b>	<b>10</b>
<b>2.1. Introduction .....</b>	<b>10</b>
2.1.1. Information, Lack of information, Misinformation.....	10
2.1.2. Science Knowledge.....	12
2.1.3. Age, Knowledge, and Don't Know Answers .....	13
<b>2.2. Methodology.....</b>	<b>15</b>
2.2.1. The Wellcome Trust Monitor Surveys.....	15
2.2.2. Measures .....	16
2.2.3. Research Questions .....	20
2.2.4. Analytical Approach .....	20
2.2.4.1. Multi-group Latent Class Analysis (LCA).....	21
2.2.4.2. Model Fit.....	23
<b>2.3. Results .....</b>	<b>24</b>
<b>2.4. Sensitivity Analysis .....</b>	<b>34</b>
2.4.1 An attitudinal measure and science knowledge .....	34
2.4.2 Covariates.....	35
2.4.3 Analytical Approach .....	36
2.4.4 Results.....	36
<b>2.5. Conclusion.....</b>	<b>38</b>
<b>References.....</b>	<b>43</b>
<b>3. Educational Disparities, Biomedical Efficacy and Science Knowledge Gaps: can the Internet help us reduce these inequalities? .....</b>	<b>47</b>
<b>3.1. Introduction .....</b>	<b>47</b>
3.1.1. Biomedical Self-Efficacy .....	50
3.1.2. The Internet, Biomedical Self-Efficacy, and Science Knowledge .....	53
3.1.3. Testing the Knowledge Gap Hypothesis .....	57
<b>3.2. Method.....</b>	<b>58</b>
<b>3.3 Results .....</b>	<b>67</b>
<b>3.4. Conclusion.....</b>	<b>74</b>
<b>3.5. Appendix.....</b>	<b>77</b>
<b>References.....</b>	<b>84</b>
<b>4. “Excuse me, doctor...” What drives people’s confidence to participate in medical decision-making? .....</b>	<b>90</b>
<b>4.1. Introduction .....</b>	<b>90</b>
4.1.1. The doctor-patient relationship .....	90
4.1.3. Preferences for participation and socio-economic differences.....	94
4.1.4. Efficacy, Trust, and Participation in medical decision-making .....	96
<b>4.2. Method.....</b>	<b>101</b>



4.2.1. Data .....	101
4.2.2. Measures .....	101
4.2.3. Analytical Strategy .....	106
<b>4.3. Results .....</b>	<b>107</b>
<b>4.4. Conclusion.....</b>	<b>114</b>
<b>4.5 Appendix.....</b>	<b>117</b>
<b>References.....</b>	<b>130</b>
<b>5. Overall Conclusion .....</b>	<b>136</b>
<b>References.....</b>	<b>143</b>

# **1. Overall Introduction**

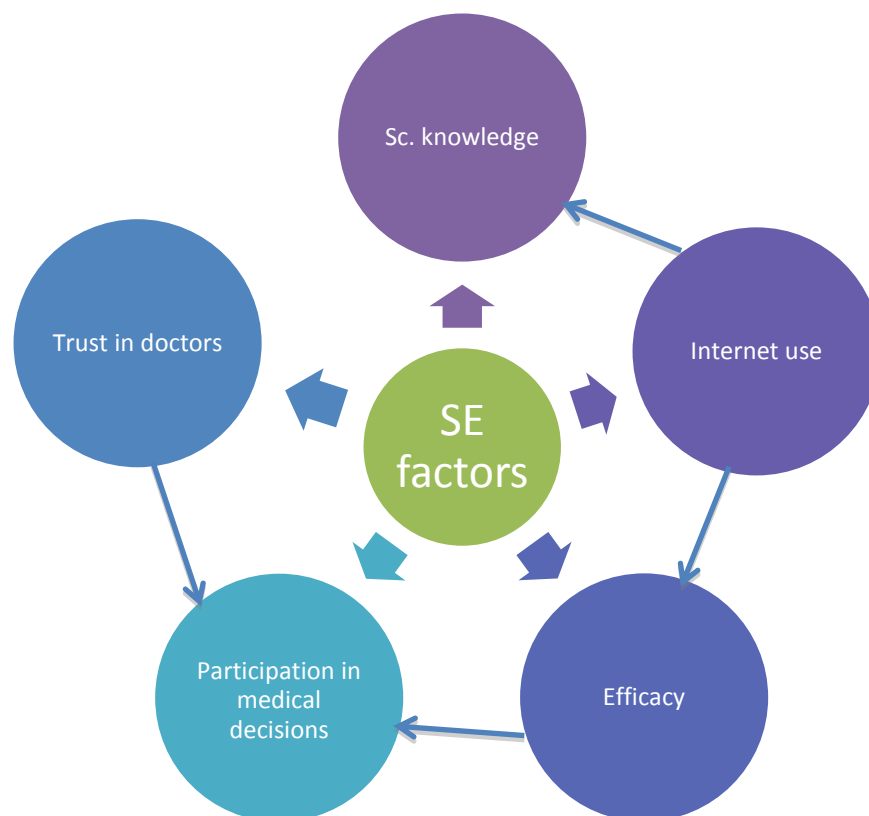
## **1.1 Thesis background**

The world is facing growing inequalities between groups with low and high socio-economic power. This poses great challenges to democracies and paves the way for even wider gaps between the rich and the poor in the future. If left unaddressed, the cycle will continue to perpetuate itself indefinitely because poverty breeds even more poverty. Information and access to information disparities have been identified as primordial in addressing this threat to the legitimacy of the democratic process by the World Economic Forum (Bilbao-Osorio et al., 2011).

Science and medicine, democratic institutions in their own right, constitute a central part in achieving progress. They are also two of the conveyors of inequality, in part due to their very specialised nature, and in part due to them being very hierarchical and traditionally paternalistic institutions. Differences in levels of scientific literacy between low and high socio-economic groups are well documented in the literature (Miller, 2001; Tichenor et al., 1970). The same holds true for public participation in the medical decision-making process (DeWalt, 2007). Addressing how we measure central constructs, such as scientific literacy and socio-economic disparities, as well as the factors that contribute to the creation and propagation of these inequalities is of vital importance in addressing the democratic deficit in information. Moreover, understanding the factors at play in giving citizens agency to take part in the decision-making process regarding their health is imperative in an effort to transition to a more democratic form of medical practice, and move from the paternalistic model of medicine towards a shared-decision model (Charles, Gafni, Whelan, 1999).

Consequently, my PhD thesis is composed of three separate papers that centre on two main themes related to the above. The first theme addresses a methodological issue, namely the measurement of science knowledge and the extent of age-related bias. The second theme tackles the complexities of the interrelationships among socio-economic factors, science knowledge, biomedical self-efficacy, and trust in doctors. **Chapter 2** investigates the extent of age-related bias in how we have traditionally measured science knowledge. Mass media effects on knowledge and efficacy gaps caused by educational differences are explored in **Chapter 3**. The role of efficacy and trust, along with socio-economic factors, in shaping public attitudes toward participation in medical decision-making is explored in **Chapter 4**. I use the Wellcome Trust Monitor Surveys (Waves I, II, and III) in all the work presented in this thesis.

A conceptual model of the relationships explored in the whole thesis is presented in the following diagram:



The conceptual model presented in the diagram only serves as an illustration of the types of relationships explored in this thesis. However, two of the central concepts represented, biomedical self-efficacy and one socio-economic factor, namely age, are measured differently in some of the empirical chapters. The variation in measurement needs to be explained before going into detail regarding the analyses conducted in each of the chapters. Firstly, biomedical self-efficacy is measured differently in Chapters 3 and 4. The sole reason for this analytical decision was a lack of data availability. In Chapter 3 I use Wave II of the Wellcome Trust Survey, whereas in Chapter 4 I use Wave III of the same survey. This decision was made because the outcome investigated in Chapter 4, namely participation in medical decision-making, was not available at Wave II.

Furthermore, the efficacy measure used in Chapter 3 is composed of two elements: 1. self-reported familiarity with biomedical terms; 2. participation in debates regarding regulating medical research; The efficacy measure in Chapter 4 is also composed of element 1, but element 2 is concerned with a different form of participation, namely discussions of science stories heard on the news. The first measure of efficacy relates to biomedicine as a more external entity. The second measure can be argued to personalise it to an extent, seeing as the participatory component is about everyday discussions of science, not public debates. There is one main implication for the conceptual model. Any positive association between Internet use and efficacy in Chapter 3 cannot be extended when discussing the impact of increased self-efficacy on participation in medical-decision making. In other words, if efficacy is found to have a positive impact on participation in medical decision-making, it would not be conceptually and empirically rigorous to say that the Internet is a medium through which that type of efficacy can be increased, unless separately tested.

Lastly, age is recoded into different categories in each chapter. The rationale for this analytical decision is twofold. Firstly, it serves a different function in Chapter 2, compared to Chapters 3 and 4. Namely, the analysis in Chapter 2 focusses on potential age-related bias in the knowledge battery at Waves I and II of the Wellcome Trust Survey. Due to data constraint and computationally demanding models, which require reasonable sample sizes, age was split into three categories: ‘14-18’, ‘19-49’, ‘50 and over’. In Chapter 3, age is measured as a continuous variable, it is not a central measure, and it is included in analyses because it is a potential confounder. In Chapter 4, age is just one of the socio-economic factors explored in relation to the outcome. It was recoded into three categories: 18-34, 35-54, 55 and over. The main implication for the conceptual model is that any relationships between age and knowledge, age and efficacy, and age and participation in medical decision-making, will carry different meanings. For instance, if age is found to be a driver of participation in medical decision-making in Chapter 4, and age is also found to have a positive relationship with efficacy in Chapter 3, any policy implications for increasing population-level confidence and participation derived from the thesis as a whole should bear in mind the different sub-populations used in each analysis.

## **1.2 Breakdown of chapters**

One of the socio-economic factors found to have an association with science knowledge is age. Older people have usually been found to know less about science than young people. This has been found using the summed score approach whereby incorrect answers and don’t knows are coupled together as wrong answers, and correct and wrong responses are added up into a scale representing an incremental measure of knowledge. Moreover, there is evidence that information, lack of information, and misinformation would have a different impact on attitudes, compared to the two-dimensional operationalisation of the construct

(information versus lack of information) that has been the focus of most Public Understanding of Science studies (henceforth 'PUS') (Luskin and Bullock, 2011).

In **Chapter 2** I explore the methodological question of whether older respondents really know less about science, or whether the questions used in the Wellcome Trust Monitor Surveys I and II to measure it behave differently in different age groups (14-18, 19-49, 50+). Multi-group Latent Class Models are employed to this end. In other words, the general question I pose is: Are older respondents really less science literate or do these questions trigger different probabilities of a correct response for people in different age brackets? In addition, I also test to see if there is evidence of misinformation. No evidence of misinformation is found. Evidence is found that the science knowledge questions work differently across the age groups. To put it simply, older people don't really know less about science (from what we can know so far), instead the questions we are using to test science literacy appear to favour younger people. A sensitivity analysis is therefore conducted in order to ascertain whether using the latent class framework accounting for age-related bias instead of the summed scale would make a practical difference when analysing a substantive matter, namely attitudes to genetics. No substantive differences are found.

In **Chapter 3**, Professor Allum and I test for science knowledge and internal biomedical self-efficacy gaps between groups of people with low and high educational backgrounds, using Wave II of the Wellcome Trust Monitor Surveys. We also test whether Internet use and paying attention to medical stories online help reduce educational disparities in knowledge and efficacy. Biomedical self-efficacy is a concept that is new to Public Understanding of Science (henceforth mostly referred to as biomedical self-efficacy or, simply, efficacy). This is mostly based on the concept of internal political efficacy (Tedesco, 2011). It taps into self-reported knowledge of biomedical terms and willingness to participate in the debates regarding regulating medical research. Since using a latent variable

framework accounting for age-related bias made no substantive difference to the attitudinal question investigated, and age differences are not the focus of this chapter, a summed science knowledge measure is employed. We model two sets of ordinary least squares regressions with three-way interactions amongst education level, internet use, and attention to medical stories online predicting, in turn, science knowledge and biomedical self-efficacy. Evidence of knowledge and efficacy gaps between low and high education groups is found. Using the Internet and paying attention to stories online are found to help reduce those differences, and thus potentially reduce the democratic deficit in information.

In **Chapter 4** I examine how two seemingly opposed factors, namely biomedical self-efficacy and deference to medical authority (in the form of trust in doctors as information sources) shape the public's confidence to challenge and query their General Practitioner's (GP) decisions regarding their health, using Wave III of the Wellcome Monitor Surveys. Very few studies employing representative samples have been carried out studying this issue, and those few were set in the US. Biomedical self-efficacy is measured slightly differently in this chapter because of the unavailability of a question on participation in the debates about regulating medical research used to construct the efficacy measure used in **Chapter 3**. Instead, questions around people's confidence to discuss science stories heard or seen on the news are used to tap into a more participatory aspect, as opposed to looking at just self-reported familiarity with biomedical terms, a common concept employed in PUS studies. Age, education, ethnicity, and gender are the socio-economic factors which are explored here. Moreover, knowing a medical professional and having a long-term illness or disability are also investigated. A conditional inference tree estimating an ordinal regression predicting willingness and confidence to query GPs is modelled. Biomedical efficacy and deference to medical authority are both included in the same model and the algorithm is

allowed to find the combinations of the two constructs that best predict confidence to participate. People who are efficacious are found to be more willing and confident to participate in decisions regarding their health, and thus subscribe to the shared model of patient-physician interaction, while those who show high deference to medical authority seem to adhere to the traditional paternalistic model of medicine.

### 1.2.1 Publication plan:

- **Chapter 2** is currently under review for publication in the *International Journal of Social Research Methodology*.
- **Chapter 3** is co-authored with Professor Nick Allum and will soon be submitted to *Science Communication*.
- **Chapter 4** has been submitted to *Social Science and Medicine*.

## 1.3 Collaboration with the Wellcome Trust

My PhD was funded through a 1+3 scholarship awarded by the Economic and Social Research Council in collaboration with the Wellcome Trust, set up by Professor Nick Allum and Dr Hilary Leever. My collaboration with the Wellcome Trust extended beyond that, however, and I will always be grateful for that opportunity. Firstly, for the first three years of my scholarship and first two years of my PhD, Dr Hilary Leever, Head of Education & Learning at Wellcome, was my formal secondary supervisor. Her guidance has been invaluable and is greatly appreciated. Secondly, every year I spent two or three weeks at the Wellcome Trust main offices in London. I have been fortunate enough to be a part of three teams: the Education & Learning Team, the Evaluation Team, as well as the Insight & Analysis Team. I have been responsible for checking all the analyses done for the report



accompanying the release of Wave II; I have advised on matters related to the design of both Wave II and III; I have conducted preliminary research on gender differences in attitudes to science as well as science capital in adults. This was then shared with a group at King's College who were working on measures of science capital in young people. I have given quite a few internal talks there, and each talk has provided much food for thought and consisted of interesting debates. The diverse nature of the teams I have been a part of has provided a great opportunity for me to discuss my research with people who have very different backgrounds to mine: specialists in the natural sciences, neuroscientists, chemists, developmental psychologists, policy experts, to name just a few. This has provided a rich experience, beyond what one can get in a single academic department and, I believe it has made me a better researcher by giving me a broader perspective on the issues I investigate in this thesis.

References cited in the 'Overall Introduction' are all to be found in the three 'References' sections appended to each empirical Chapter henceforth.



## **2. Information, Misinformation, Lack of Information, and Age-based Differences in Science Knowledge: A Multi-group Latent Class Model.**

### **2.1. Introduction**

#### 2.1.1. Information, Lack of information, Misinformation

The role of information in shaping public attitudes to a wide range of socio-political issues, as well as social behaviour, is well documented in the public opinion literature. The importance of citizens being factually informed has been the unwavering normative thrust in public opinion research since the publication of Converse's classic study in 1964. The study of how people's misperceptions affect their attitudes has lain primarily within the realms of experimental psychology and political science. The latter has perhaps been more filled with applications of experimental research into its main areas of interest, such as attitudes to political issues, political candidates or incumbents, as well as voting behaviour (Lewandowsky, Ecker, Seifert, Schwarz, and Cook, 2012).

Scholars have focused on the role of ignorance and the various ways it influences public opinion. In other words, they have focused on distinguishing citizens who are informed from those who are uninformed, and the consequences this might have on their views and behaviour. However, a significant advance was made by Kuklinski, Quirk, Jerit, Schwieder, and Rich (2001), who posited that there are fundamental notional differences between simply being uninformed about a fact and being confidently wrong about it, or being misinformed (See also Luskin and Bullock, 2011). They put forward a conceptualisation of

three terms: information (correct answers), lack of information (saying ‘don’t know’), and misinformation (thinking you know something and getting it wrong). Moreover, since knowledge is advocated as one of the pillars of a democracy, scholars have argued that there is a societal cost of misinformation. Thus, if a majority accepts as true a fact that is incorrect, this misperception may form the basis for political and societal decisions that offset the best interest of a democracy. In a similar vein, it is argued that if individuals are simply wrong-headed about something, they may make choices, for themselves and their families, that are not in the least beneficial and can come at a great cost (Kuklinski et al., 2001).

Kuklinski et al. (2001) postulate that dependence on misinformation is different from mere ignorance, defined as the absence of knowledge. Ignorance may be less pernicious than misinformation because, in the absence of knowledge, people often turn to simpler heuristics when making decisions. These can work well, at least under the right circumstances. For example, simple awareness of an object often allows citizens to make acceptable deductions about it. Also, on average, people tend to not rely a great deal on decisions made solely on the basis of such heuristics (Dancey and Sheagley, 2013). Lack of knowledge about a fact or a set of facts hardly ever drives support for a cause, in contrast to false beliefs based on misinformation, which are often held strongly and with conviction. For instance, studies have shown that those who most heartily snub the scientific evidence for climate change are also those who believe they are best informed about the subject (Lewandowsky et al., 2012). Thus, the measures developed here, aiming to distinguish between these three types of respondents, could potentially help us to better comprehend the mechanisms underpinning the relationship between information and attitudes to controversial issues, such as vaccinations.

### 2.1.2. Science Knowledge

Capturing the construct of science knowledge has been a central concern in PUS research from the very beginning. Social surveys have been an essential tool for researching science knowledge and attitudes towards science and technology. The first survey used in PUS was Withey's in 1957, in the United States (Stares, 2008; Sturgis and Allum, 2004). Another important study was the US National Science Board's biennial Science Indicators Survey in 1979. The first survey of attitudes towards science conducted in the UK was Durant et al. in the late 1980s (Durant, Evans, and Thomas, 1989), and there are quite a few still ongoing. There appear to be two surveys that are either completely dedicated to topics pertaining to PUS, or just contain one or several PUS topic(s): the British Social Attitudes series, sponsored by the Wellcome Trust, government bodies and government-commissioned bodies (Stares, 2008), and the Wellcome Trust Monitor Surveys, with two waves completed. The latter is used in the present study. These surveys include items/questions designed to tap into people's information levels, lack thereof, and, potentially, misinformation.

Science knowledge is posited as a multidimensional concept. In the US, Jon D. Miller (1983, 1992, 1998, 2004) argued that it has three linked dimensions, namely 'a vocabulary of basic scientific constructs sufficient to read competing views in a newspaper or magazine [...], an understanding of the process or nature of the scientific inquiry [...], some level of understanding of the impact of science and technology on individuals and on society.' UK scholars have put forward a fourth dimension, though, postulating that leaving science actors out of the discussion would lead to an insufficient comprehension of the scientific process (Bauer, Durant, Evans, 1994; Bauer, Petkova, Boyadjieva, 2000).

The measures of science knowledge analysed here have a focus on biomedicine, particularly genetics. The most employed method of constructing the continuous measure of science

knowledge is the sum-score method mentioned above (adding up all the correct responses). The concern arising from utilising this method is at least threefold. Simply creating a score from all the correct answers omits taking into account measurement error. Second, all items are assumed to be of the same difficulty. Third, ‘don’t know’ answers are coupled with outright incorrect answers, where there is an apparent substantive difference between people admitting they do not have the information and those who answer incorrectly. This could be either because they think they know but they don’t, or due to a social desirability bias.

### 2.1.3. Age, Knowledge, and Don’t Know Answers

The Wellcome Trust Monitor Survey is fielded to two samples, one of adults (aged 18+), and one of young people (aged 14-18). The survey contains multiple items designed to assess science knowledge and engagement with biomedical science. Analysis of the first Monitor Survey yielded a curvilinear relationship between age and high and low levels of scientific knowledge. High scorers were lowest among those aged 65 years and over. The proportion of high scorers for the middle age groups was around three in ten (of those aged 35-49 and 50-64). For low scorers, the reverse pattern holds true, and for young people, 14-16 year olds were less likely to be high scorers than 17-18 year olds. However, these observations cannot be taken unproblematically to indicate true differences in knowledge about science between these age groups. Another possibility is that a given knowledge item has a different probability of generating a correct response from different age groups even if each age group has the same underlying level of knowledge. The cut-off point for the older age group examined in this study is 50 because it still makes substantive sense and it provides an adequate sample size for the multi-group analyses (Butt, Abeywardana, and Phillips, 2009).

Converse (1964) stated that if a respondent is presented with an attitude/opinion item regarding an unfamiliar issue, they are very likely to say they don’t know. Sicinsky (1970)

argues that this holds true for knowledge batteries as well as opinion batteries. There have been some debates in political science around the issue of what ‘don’t know’ answers represent, i.e. hidden knowledge, propensity to guess, etc. Luskin and Bullock (2011) show that, in the case of political knowledge at least, these ‘don’t know’ answers are not covering hidden knowledge. People who say they don’t know really don’t know. Political studies have also shown older respondents have a higher chance of saying they don’t know than younger individuals. They also found that, in regression analyses, when controlling for education, gender, and race, that relationship tends to disappear (Ferber, 1966; Gergen and Back, 1965; Glenn, 1966; Francis and Busch, 1975). Be that as it may, when comparing age groups or merely individuals of different ages in terms of their knowledge levels, one can draw erroneous conclusions about an older respondent’s level of knowledge/ignorance/misinformation when not taking into account any potential measurement effects (here understood as possible non-equivalence of items across three age groups). In other words, these observations about the relationship between ‘don’t know’ answers and age cannot be accepted for what they are without investigating whether they are indeed true or an artefact of measurement.

This study has two primary goals: 1) to unearth information, lack of information, and misinformation measures from the data; 2) examine the extent to which the ten science knowledge questions work the same across the three age groups using a multi-group latent class model (Lazarsfeld and Henry, 1968; Clogg and Goodman, 1984). The latter is formally called measurement invariance or homogeneity, a term more specific to latent class terminology. Measurement invariance across groups/populations is presumed to characterise a scale/test/quiz when respondents with the same score on the underlying (latent) concept have the same expected raw/true score at the item level, the subscale total score level, or both. Essentially, the question is whether the test performs the same for each group of

respondents. Lack of measurement equivalence would render significant comparisons of observed mean score differences problematic. If the condition of measurement invariance does not hold, the aforementioned differences could indicate both the true difference between groups and a disparity between the underlying construct and the measured score that varies across groups. If, however, the condition of measurement invariance does hold, the relationship between the latent variable and each manifest variable remains invariant across groups. Thus, any observed difference can be taken as being indicative of only true differences between groups (Byrne, Shavelson, and Muthén, 1989; Byrne and Watkins, 2003; Drasgow, Levine, and Williams, 1985).

## **2.2. Methodology**

### 2.2.1. The Wellcome Trust Monitor Surveys

The data used in this study come from the Wellcome Monitor Survey Waves I and II. The Wellcome Monitor is a triennial survey of adults and young people in the UK, based on a representative probability sample. Its main purpose is to provide a more accurate and nuanced picture of interest in, knowledge of, and attitudes to medical research, with a focus on biomedicine. Wellcome Monitor I took place in 2009 and interviews were conducted with 1179 adults (aged 18+) and 374 young people (14-18 years old; 121 part of a ‘core’ sample, and 253 part of a ‘boost’ sample). The response rate was 49.3% for the adult sample, 86.4% for the young ‘core’ sample, and 64.2% for the ‘boost’ sample (Butt et al., 2009). A complete report of the sampling procedure as well as other aspects of the survey design can be found in Butt et al. (2009). Wellcome Monitor II took place in 2012 and interviews were conducted with 1,396 adults (aged 18+) and 460 young people (14-18 years old; 154 part of a ‘core’ sample, and 306 part of the ‘boost’ sample). The response rates for Monitor II were



52.6% for the adult sample, 75.4 % for the young ‘core’ sample, and 64.4% for the ‘boost sample’ (Clemence et al., 2013).

All data preparation and analyses were carried out in Stata 12, SPSS 19, MPlus 7.31 (Muthén & Muthén, 1998-2010), and R 3.2.0/RStudio 0.99.893. More specifically, the two cross-sectional data sets for Waves I and II of the Wellcome Monitor Survey (2009 and 2012, respectively) were combined using the append command in Stata 12. The final sample size is fairly large, with a total of 3,409 respondents. Furthermore, all the recode procedures were carried out in SPSS, the latent class models were specified in MPlus 7.13, and the results were then exported to R for post-processing purposes. This was motivated by the fact that for latent class models with nominal indicators MPlus only gives results on a logit scale. These were then converted to probabilities in R. An additional model fit test, the likelihood ratio test, is also performed in R as it is beyond MPlus’s capabilities (Kuha, 2013).

### 2.2.2. Measures

#### *The Science Knowledge Items*

Ten items measuring knowledge of science (with a focus on genetics) have been asked of respondents from both Wave I and II. Moreover, nine of the ten items analysed here are ‘textbook’, factual, quiz questions, and a tenth is designed to gauge the respondents’ understanding (or lack thereof) of controlled experiments. Nine statements were shown to the respondents and they were asked to choose one of five response options. The correct answer is given in brackets.

- 1) Electrons are smaller than atoms. (T)
- 2) All radioactivity is man-made. (F)
- 3) All plants and animals have DNA. (T)

- 4) More than half of human genes are identical to those of mice. (T)
- 5) The cloning of living things produces identical copies. (T)
- 6) The oxygen we breathe comes from plants. (T)
- 7) Lasers work by focusing sound waves. (F)
- 8) By eating a genetically-modified fruit, a person's genes could also become modified. (F)
- 9) It is the mother's genes that determine the sex of the child. (F)

Respondents were presented with these nine statements and were then asked to choose one of the following response options:

- 1) Definitely true
- 2) Probably true
- 3) Probably false
- 4) Definitely false
- 5) Don't know

Item 10 measures the respondents' understanding of controlled experiments. The item is worded as follows: 'Suppose a drug used to treat high blood pressure is suspected of having no effect. On this card, there are 3 different ways scientists might use to investigate the problem. Which one do you think scientists would be likely to use?' Respondents are then asked to select one of the following options:

- 1) Talk to those patients that have used the drug to get their opinion?
- 2) Use their knowledge of medicine to decide how good the drug is?
- 3) Give the drug to some patients, but not to others, then compare the results for each group?
- 4) Don't know.

The ten knowledge items were recoded so that each new item had three nominal categories denoting correct, incorrect, and ‘don’t know’ responses, each meant to represent information, misinformation, and lack of information, respectively. I regard all incorrect answers as indicating a degree of misinformation, whether respondents say ‘probably’ or ‘definitely’ when answering wrongly. ‘Probably’ indicates some degree of confidence, and even though wrong beliefs held in this manner are probably less pernicious than those held with ultimate confidence, I posit that it would still make it very likely for someone to act on their wrongly held beliefs about an issue, as opposed to people who just admit they ‘don’t know’. On a different note, however, those saying ‘probably’ are most likely less resistant to change than individuals who exhibit more confidence in their answer. However, this particular nuanced distinction is for another study to determine. Table 2.0 presents frequencies for the three-category nominal science knowledge items.

**Table 2.0 Frequencies for knowledge items (%)**

<b>Variable</b>	<b>Correct</b>	<b>Incorrect</b>	<b>Don’t know</b>
<b>Electrons</b>	46.2	22.2	31.6
<b>Radioactivity</b>	65.2	19.8	15.0
<b>DNA</b>	88.5	5.1	6.4
<b>Mice</b>	47.7	23.1	29.2
<b>Cloning</b>	74.5	11.6	13.8
<b>Plants</b>	82.6	11.8	5.7
<b>Lasers</b>	44.0	24.5	31.6
<b>Genetically Modified</b>	69.4	11.2	19.4
<b>Mother’s genes</b>	63.5	19.0	17.5
<b>Controlled experiments</b>	64.4	31.1	4.5

### *Age*

Age is measured as a continuous variable. For the purpose of the multiple group analysis, it was recoded as a categorical variable, yielding three age groups: 14-18 (863 respondents/25.3%), 19-49 (1,233 respondents/36.2%), and over 50 (1,310 individuals/38.5%). The cut-off points for the categories were informed by the data and model type. Multi-group latent

class models are sensitive to sample sizes within each group. This breakdown provided a large enough sample within group for me to be able to proceed with multi-group modelling and test for measurement invariance. As a result, the groups were formed as such so that they would make some substantive sense, as well as provide reasonable sample sizes for the models to converge successfully. Future analyses employing larger samples should endeavour to split the middle group into two groups for a more nuanced analysis. The final number of observations retained by Mplus following listwise deletion is 3,124 individuals, down from 3,409. The Maximum Likelihood algorithm uses all the information available. Moreover, it assumes that the data is missing at random.

A complete case analysis was preferred to one which allowed for item missingness because the objective of this study is to measure underlying science knowledge using all of the knowledge items in the two waves of the survey. In addition, age-related measurement invariance with regard to each of the items is investigated. In order to be able to make significant comparisons across all items, one has to examine the same individuals, who have data on all items. If different individuals with data on only a subset of the items are used, comparisons for each item are invalid because those individuals' underlying level of knowledge is given by what is essentially a different scale, composed of only a subset of items. It would render judgements about invariance in relation to each item, and across all items, in the battery invalid. One limitation is that the discarded data might not be missing at random, and item non-response has introduced bias in assessing the overall level of knowledge. This study should only be regarded as an exercise in testing for measurement invariance in relation to this specific battery of knowledge items. Findings should only be interpreted in relation to the remaining sample and its associated characteristics presented in Table 2.0. No claims of representativeness are made.

### 2.2.3. Research Questions

In the absence of much previous research on age-related bias in science knowledge scales, three research questions were set instead of formulating hypotheses.

**Q1:** How many latent classes emerge from the data? Do the latent classes appear to be ordered in terms of their conditional probabilities? If yes, how many underlying dimensions can be inferred from the latent classes derived from the data?

**Q2:** What is the extent of any potential age-related bias? In other words, are the items invariant across the three age groups? Who are these items better or worse for? Are older respondents more likely to give a ‘don’t know’ response?

**Q3:** For future Wellcome Trust Monitor Surveys, what recommendations can be made for the design of this battery of questions?

### 2.2.4. Analytical Approach

Science knowledge cannot be directly observed, which is why it is called a latent variable. What we do observe are the ten items, called indicator/manifest variables, intended to encapsulate this construct. By examining the pattern of associations amongst these manifest variables we can infer there is an underlying latent variable (or more; for instance, to distinguish between an information/misinformation dimension and one of ‘don’t know’/lack of information) (Bollen, 1989; Bartholomew, Knott, and Moustaki, 2011; Skrondal and Rabe-Hesketh, 2004). The purpose of a measurement model is to describe how well the indicator/manifest variables serve as a measurement instrument for the latent variables. Moreover, they allow for measurement error, and provide indications for potential ways in which the measurements can be improved.

### 2.2.4.1. Multi-group Latent Class Analysis (LCA)

All items are treated as nominal. No a priori assumption about the ordering of categories is made. A multi-group latent class model can be estimated using three different types of parameterisation: logistic, probabilistic, and log-linear. MPlus estimates it as a multinomial logistic regression model with ten nominal observed variables as independent variables, and a categorical latent variable as the dependent variable. The goal is to categorise respondents into latent classes using the manifest indicators that best differentiate between classes (Nylund, Asparouhov, and Muthén, 2007). Maximum Likelihood estimation was used for all models. The multi-group models were estimated using age group as a covariate, as it is a more flexible approach. The results are then post-processed in R and the logits obtained in MPlus are converted into probabilities, as their interpretation is far more intuitive (Kuha, 2013). The basic multi-group latent class measurement model can take the following form:

$\eta$  is a categorical latent variable with  $c$  unordered categories or classes ( $c=1, \dots, C$ )

$y_j$  are categorical observed indicators ( $y_j = 1, \dots, p$ ) with  $L_j$  possible levels ( $l=1, \dots, L_j$ )

$g$  is a categorical observed variable that is treated as a predictor of  $\eta$

$$\eta_c = P(\eta=c), c=1, \dots, C \quad (1)$$

$$\pi_{jl}(c) = P(y_j=l | \eta = c) \quad (2)$$

We are interested in modelling two types of probabilities:

- Prior probabilities of belonging to class  $c$  (1)
- Conditional response probabilities (2) – the probability of responding in category  $l$  to item  $j$ , given membership of latent class  $c$

In order to investigate measurement invariance I add age groups (**g**) as covariates. These are specified as three dummy variables, with the older group as the reference group. Conditional probabilities help us reach a substantive interpretation of the classes by examining the patterns of responses. For instance, if respondents in a class have very high probabilities of saying they don't know, then we can label this as a 'don't know' or 'lack of information' class. Class probabilities tell us the proportion of people that fall into a particular class **c**.

After inspecting class and conditional response probabilities for the model estimated on the pooled sample, a latent class model is estimated in each of the three age groups to find the model that best fits the data in each group separately. Once construct equivalence has been established (whether the classes have similar interpretations in all groups), we can proceed to investigate measurement invariance. The multi-group analyses start with a model where the measurement model is held equal across groups. Since MPlus uses the logistic parameterisation, intercepts and loadings are fixed to be equal in all groups. This analysis is exploratory by nature, so no part of the measurement model was fixed to specific values. Next, a model where everything is free to vary (unconstrained) is estimated and compared to the fully constrained model. If it is deemed to be better, the next step is to examine each item in turn to see where the misfit stems from. Moreover, intercepts, and then intercepts and loadings, are freed for each item in turn, examining fit statistics to see if the model fits better than the fully constrained model. By freeing intercepts, we model a direct effect between age group and response probabilities to any given item. In layman's terms, an older respondent might be more likely than a younger respondent to say 'don't know' to an item, regardless of class membership. Only freeing a slope/intercept allows the relationship between an item and the categorical latent variable to vary by group. By freeing both intercepts and loadings/slopes we allow an interaction between latent class, group, and response probabilities to an item. What this allows is the effect of age group on response probabilities

to items to vary across the latent classes. The item could have a dissimilar interpretation in relation to the latent variable in different groups (McCutcheon, 1987).

#### 2.2.4.2. Model Fit

How well a model fits the data can be assessed in a number of ways. I used the Akaike Information Criterion (AIC), the sample-adjusted Bayesian Information Criterion (BIC), the likelihood ratio test, and the bootstrapped likelihood ratio test (BLRT) (Asparouhov and Muthén, 2012; Kuha, 2013; McCutcheon, 1987; Hagenaars and McCutcheon, 2002; Kankaras and Moors, 2009, Nylund et al., 2007). The AIC and BIC formulas used by Mplus are:

$$\text{AIC} = -2\log L + 2*r \quad (3)$$

where  $\log L$  is the log-likelihood value, and  $r$  is the number of free parameters;

$$\text{BIC} = 2\log L + p*\ln(n) \quad (4)$$

Where  $p$  is the number of parameters and  $n$  is the sample size;

Various models are compared using these indices, with the smallest value on any given index denoting the model that fits the data best. The two indices do not always agree, however. The BIC is a better indicator of the number of classes than the AIC, so it will receive more attention (Nylund et al., 2007).

The likelihood ratio test for nested models compares the log-likelihood of two nested models, and it provides an LR test statistic, which is referred to the chi-square distribution with a number of degrees of freedom (the difference in the number of parameters estimated in the two models). The LR statistic has an associated p-value (if  $<0.05$  then the first model is rejected in favour of the second model); this test is not available in MPlus so it is conducted in R using the 'lcat' post-processing functions. It is used to compare more restricted models with less restricted models, such as when a parameter is free (Kuha, 2013).



The bootstrapped likelihood ratio test (BLRT) is used to select the appropriate number of classes; it compares a model with  $k-1$  classes to a model with  $k$  classes. It achieves this by computing a test statistic based on the  $-2 \times \log(\text{likelihood difference})$  between the two models; it also has an associated p-value which helps decide if the  $k-1$  classes model is rejected in favour of the  $k$  classes model.

LCA was chosen for two primary reasons. First, the normal linear factor model is inappropriate when dealing with categorical manifest variables, and the reasons for that have been widely discussed in the literature (Bartholomew et al., 2011). Second, since science knowledge is postulated to be a continuous measure, latent trait models were deemed appropriate and were estimated first, but failed to converge. Increasing the number of starts until a maximum of 5000 was reached, increasing the number of iterations, setting *stscale* to 1, or building the model in blocks by estimating separate models in the three groups to only bring it all together in the last step, as per the Mplus User Guide (Muthén and Muthén, 1998-2010), were all futile attempts. With a latent class model we can still infer that there is an underlying continuous measure if the latent classes appear to be ordered in terms of their conditional response probabilities, even though the latent variable itself is not ordinal in nature. Finally, multi-group LCA offers considerably more flexibility than latent trait analysis, the factor model, or item response theory models, in which stronger distributional assumptions have to be satisfied (Kankaras and Moors, 2009).

### **2.3. Results**

An analysis of the pooled data reveals a common 4-class structure in terms of the BIC statistic, and five according to the BLRT (Table 2.1). Separate analyses of the three age groups reveal a 3-class model to best fit the data for the young group in terms of the BIC and the BLRT. For the middle group a three-class solution also appears to fit the data better

according to all fit diagnostic statistics, whereas a 4-class structure emerges for the third group, aged 50+, according to all fit statistics. Table 2.2 provides fit statistics for all analyses conducted separately in the three groups.

**Table 2.1 Fit statistics for the pooled sample**

<b>Model</b>	<b>BLRT</b>	<b>p(bootstrapped)</b>	<b>AIC</b>	<b>BIC</b>
<i>2 classes</i>	3179	<0.001	53602	53853
<i>3 classes</i>	611	<0.001	53033	53413
<i>4 classes</i>	464	<0.001	52611	53120
<i>5 classes</i>	114	<0.001	52539	53177
<i>6 classes*</i>				

\*model not identified

The three classes also appear to be ordered in terms of their conditional probabilities, ranging from the least knowledgeable in class 1, with individuals in this class having very high probabilities of saying they don't know the answer (and so professing ignorance; henceforth the 'don't know class'). The common denominator for the second and third classes is the respondents' propensity to offer a substantive response to the question. Moreover, individuals in the second class have mixed probabilities of offering correct and wrong answers, the former being generally higher than the latter (henceforth the 'medium knowledge class').

The same applies to the third class (henceforth the 'high knowledge class'), except the probabilities of offering substantively correct answers are higher than in the medium knowledge class. For older respondents, aged 50 and over, a 4-class model fits the data better. The fourth class would substantively fit between the first and second class of the three-class structure model that fits in all age groups (henceforth the 'low-medium knowledge class').

In other words, some older respondents fall into an additional class where they offer a

**Table 2.2 Model fit statistics for each group taken separately**

<b>Model</b>	<b>BLRT</b>	<b>p(bootstrapped)</b>	<b>AIC</b>	<b>BIC</b>
<b>Young (14-18 years old)</b>				
<i>2 classes</i>	540	<0.001	13328	13523
<i>3 classes</i>	155	<0.001	13214	13509
<i>4 classes*</i>	151	0.24	13105	13500
<i>5 classes</i>	65	0.67	13082	13577
<b>Middle (19-49 years old)</b>				
<i>2 classes</i>	1130	<0.001	18512	18722
<i>3 classes</i>	190	<0.001	18365	18682
<i>4 classes</i>	139	0.38	18268	18692
<i>5 classes</i>	68	<0.001	18242	18774
<b>Older (50+ years old)</b>				
<i>2 classes</i>	1505	<0.001	21117	21329
<i>3 classes</i>	377	<0.001	20781	21103
<i>4 classes</i>	198	<0.001	20626	21056
<i>5 classes</i>	64	<0.001	20604	21143

\*model not identified

substantive response, and not just an outright ‘don’t know’. However, their probabilities of offering a correct answer are much lower than those of individuals in the medium knowledge class for the 3-class solution. Thus, it appears that they are a bit more spread out across classes than young respondents or those in the 19-49 age group, which is in keeping with the common structure found in the pooled sample.

Even though the model for the older group differs slightly in that it presents an additional class, resembling an intermediate knowledge level between the first and second classes of the three-class solution, construct equivalence is not violated as LCA is more flexible than other approaches, such as Latent Trait Analysis or Confirmatory Factor Analysis. LCA permits the estimation of models with more classes in one or more groups, setting the proportion to 0 in the additional class(es) for the groups where fewer classes provide the

best-fitting model (Kankaras and Moors, 2009). This was achieved using the ‘Training’ option in MPlus (Muthén and Muthén, 1998-2010). The multi-group models testing for invariance were estimated with three classes for individuals in the young and middle age groups, and four classes for the third age group. For the first two groups the proportion was set to 0 for the fourth class, meaning that respondents aged 14-18 and 19-49, respectively, were not allocated to this last class.

Prior probabilities for all three groups are shown in Table 2.3. It is apparent that there are only 0.1% more older respondents in the ‘don’t know’ class than there are middle-aged people. The proportion of older people in this class is a bit over double that of younger respondents, though. Most young respondents are in the medium knowledge class; 50% of middle-aged people fall in the most knowledgeable class, with older respondents not too far behind. It seems that a higher proportion of older respondents are allocated to the last class, compared to younger respondents. Table 2.4a presents the conditional response probabilities for the three groups. Table 2.4b shows probabilities for the 4<sup>th</sup> class fitted for older respondents. The conditional probabilities for older respondents are from the final 4-class model estimated for this group. The differences between the conditional probabilities derived from a four-class model and the conditional probabilities obtained in a three-class solution are so small that it makes no substantive difference whatsoever (mostly at the 0.01 level).

The pattern of conditional probabilities across the three groups was deemed to be similar enough to warrant proceeding to test for measurement invariance.

A few items have reverse probabilities for one or two of the age groups in some classes, however. The first item that exhibits this problem is ‘electrons’ in the medium knowledge class, where young respondents have a very high probability of giving a substantive, correct,

answer; respondents in the second and third groups have very high probabilities of saying they don't know.

**Table 2.3 Prior probabilities of belonging to a given latent class\***

	<b>Don't know</b>	<b>Medium Knowledge</b>	<b>High Knowledge</b>	<b>Low-medium Knowledge</b>
<b>Age group</b>	<b>Class#1</b>	<b>Class#2</b>	<b>Class#3</b>	<b>Class#4</b>
<b>Younger</b>	0.07	0.69	0.24	0.00
<b>Middle</b>	0.14	0.36	0.50	0.00
<b>Older</b>	0.15	0.10	0.42	0.33

\*derived from the multi-group model

A similar pattern can be observed in the don't know class for the item 'plants', and the medium knowledge class for items 'mice' and 'lasers'. For the don't know class, item 'oxygen', it is only older respondents who have high probabilities of giving a 'don't know' answer. For the medium knowledge class, item 'mice', something similar occurs, except the difference between probabilities of giving a substantive answer or saying 'don't know' is not very substantial. The same stays true for the same item in the high knowledge class, but only for the young respondents. Moreover, this also seems to apply to the older respondents for the 'mother's genes' item, the medium knowledge class, who do not have very disparate probabilities of giving a correct, incorrect, or 'don't know' answer. The same can be observed regarding young respondents' probabilities of answering correctly or incorrectly the question on controlled experiments in the high knowledge class. The difference in the conditional probabilities is not very large. It is then obvious that some of these questions do not work exactly the same in some classes for the selected three age groups (highlighted in the table). In other words, some of these items seem relatively independent across groups.

Given this finding, further investigation into the source of this seeming non-equivalence is warranted. In the first instance, the fully constrained model (where all parameters are held

equal across groups) was tested against a model where all parameters are free to vary (Table 2.5). The likelihood ratio test, the AIC, and the BIC, all reject the model where everything is held equal in favour of the fully unconstrained model. The next step was to identify the items that cause the most trouble. In other words, to ascertain the greatest source of misfit. To check this, likelihood ratio tests were conducted where each item was freed in turn. In the first instance, only intercepts were allowed to vary across the three groups for each item in turn.

Then each model, where the intercept for each item was allowed to vary across groups, was tested against the fully constrained model to see if by freeing that item there was any significant improvement in how well the model fits the data. The items are ordered going from the best to the worst fitting models, compared to the fully constrained model. If freeing the intercept for an item provides a significant improvement in fit, then we can accept that model as the better model. All fit statistics reject the constrained model (everything equal), in favour of the models where a direct effect from group to each item is allowed (Table 2.6).

Subsequently, both intercepts and slopes were allowed to vary across groups for each item in turn. This essentially means that an interaction between latent class, group, and item is allowed. Each model was then tested against the equal measurement model. It is clear from table 2.7 that the likelihood ratio test, the AIC, and the BIC, all reject the constrained model in favour of each model where both intercepts and slopes were free to vary. The items are again ordered in terms of improvement in fit, with the first one providing the biggest improvement in fit. Thus, all these questions appear to work differently for these three age groups. As a last step, for each item, the model allowing a direct effect was tested against the model that allowed a group X latent class X item interaction.

**Table 2.4a Conditional probabilities for each group separately for three classes**

		<b>Don't know</b>	<b>Medium knowledge</b>	<b>High knowledge</b>
<b>Item</b>		<b>Class#1</b>	<b>Class#2</b>	<b>Class#3</b>
		<b>Young/Middle/Older</b>	<b>Young/Middle/Older</b>	<b>Young/Middle/Older</b>
<b>k1</b>	<b>Electrons</b>			
	correct	0.11/ 0.03/ 0.00	0.81/ 0.21/ 0.13	0.55/ 0.60/ 0.55
	incorrect	0.00/ 0.04/ 0.00	0.18/ 0.13/ 0.09	0.26/ 0.30/ 0.26
	don'tknow	0.89/ 0.94/ 1.00	0.02/ 0.65/ 0.78	0.18/ 0.09/ 0.16
<b>k2</b>	<b>Radioactivity</b>			
	correct	0.05/ 0.00/ 0.02	0.89/ 0.40/ 0.37	0.59/ 0.82/ 0.95
	incorrect	0.27/ 0.10/ 0.07	0.11/ 0.27/ 0.30	0.28/ 0.18/ 0.03
	don'tknow	0.68/ 0.90/ 0.91	0.00/ 0.33/ 0.33	0.13/ 0.00/ 0.02
<b>k3</b>	<b>Plants</b>			
	correct	0.63/ 0.38/ 0.34	0.96/ 0.92/ 0.84	0.90/ 0.93/ 0.95
	incorrect	0.00/ 0.10/ 0.01	0.03/ 0.03/ 0.02	0.08/ 0.06/ 0.03
	don'tknow	0.37/ 0.60/ 0.65	0.01/ 0.04/ 0.14	0.02/ 0.01/ 0.02
<b>k4</b>	<b>Mice</b>			
	correct	0.11/ 0.11/ 0.01	0.55/ 0.33/ 0.43	0.17/ 0.63/ 0.71
	incorrect	0.10/ 0.06/ 0.02	0.30/ 0.33/ 0.12	0.48/ 0.22/ 0.09
	don'tknow	0.79/ 0.78/ 0.97	0.16/ 0.45/ 0.45	0.35/ 0.15/ 0.19
<b>k5</b>	<b>Cloning</b>			
	correct	0.17/ 0.22/ 0.12	0.92/ 0.64/ 0.72	0.65/ 0.81/ 0.90
	incorrect	0.00/ 0.00/ 0.00	0.07/ 0.10/ 0.05	0.20/ 0.15/ 0.07
	don'tknow	0.83/ 0.78/ 0.88	0.02/ 0.26/ 0.23	0.15/ 0.03/ 0.03
<b>k6</b>	<b>Oxygen</b>			
	correct	0.69/ 0.51/ 0.30	0.94/ 0.85/ 0.77	0.88/ 0.87/ 0.82
	incorrect	0.08/ 0.05/ 0.03	0.06/ 0.08/ 0.14	0.11/ 0.12/ 0.16
	don'tknow	0.23/ 0.44/ 0.67	0.00/ 0.07/ 0.10	0.02/ 0.01/ 0.03
<b>k7</b>	<b>Lasers</b>			
	correct	0.06/ 0.00/ 0.00	0.60/ 0.19/ 0.15	0.33/ 0.62/ 0.72
	incorrect	0.03/ 0.00/ 0.04	0.17/ 0.26/ 0.32	0.36/ 0.21/ 0.16
	don'tknow	0.96/ 0.94/ 0.96	0.22/ 0.55/ 0.54	0.31/ 0.18/ 0.13
<b>k8</b>	<b>GM fruit</b>			
	correct	0.05/ 0.05/ 0.08	0.86/ 0.63/ 0.47	0.61/ 0.85/ 0.86
	incorrect	0.14/ 0.01/ 0.08	0.12/ 0.10/ 0.15	0.17/ 0.10/ 0.04
	don'tknow	0.81/ 0.94/ 0.84	0.03/ 0.28/ 0.38	0.22 /0.05/ 0.10
<b>k9</b>	<b>Mother's genes</b>			
	correct	0.32/ 0.33/ 0.19	0.82/ 0.59/ 0.44	0.60/ 0.74/ 0.71
	incorrect	0.16/ 0.19/ 0.04	0.13/ 0.16/ 0.23	0.25/ 0.18/ 0.16
	don'tknow	0.52/ 0.49/ 0.77	0.05/ 0.26/ 0.34	0.15/0.08/ 0.13
<b>k10</b>	<b>Experiments</b>			
	correct	0.23/ 0.24/ 0.27	0.77 /0.55/ 0.55	0.42 /0.77/ 0.89
	incorrect	0.42/ 0.37/ 0.41	0.21 /0.40/ 0.41	0.54 /0.21/ 0.09
	don'tknow	0.36/ 0.39/ 0.31	0.02 / 0.05/ 0.04	0.05 /0.01/ 0.04

**Table 2.4b Conditional probabilities for Class#4 in the older group**

	<b>Item</b>	<b>Class#4</b>
<b>k1</b>	correct	0.44
	incorrect	0.34
	don'tknow	0.22
<b>k2</b>	<b>Radioactivity</b>	
	correct	0.43
	incorrect	0.48
<b>k3</b>	<b>Plants</b>	
	correct	0.74
	incorrect	0.23
<b>k4</b>	<b>Mice</b>	
	correct	0.53
	incorrect	0.45
<b>k5</b>	<b>Cloning</b>	
	correct	0.61
	incorrect	0.34
<b>k6</b>	<b>Oxygen</b>	
	correct	0.71
	incorrect	0.26
<b>k7</b>	<b>Lasers</b>	
	correct	0.39
	incorrect	0.53
<b>k8</b>	<b>GM fruit</b>	
	correct	0.69
	incorrect	0.24
<b>k9</b>	<b>Mother's genes</b>	
	correct	0.61
	incorrect	0.34
<b>k10</b>	<b>Experiments</b>	
	correct	0.41
	incorrect	0.56
	don'tknow	0.02



**Table 2.5 Measurement model equal versus unconstrained**

	<b>Model</b>	<b>logL</b>	<b>LR</b>	<b>df</b>	<b>p-value</b>	<b>AIC</b>	<b>BIC</b>
1	Measurement model equal	-26398	394	6	<0.001	52962	53471
2	vs Unconstrained	-26201				52580	53126

*\* logL =log likelihood; H0: measurement model equal; H1: measurement model free to vary across groups*

**Table 2.6 Model fit statistics - direct effect only for each item**

<b>Model</b>	<b>LR</b>	<b>df</b>	<b>p-value</b>	<b>AIC</b>	<b>BIC</b>
<i>Direct effect only *</i>					
<b>Electrons</b>	498	10	<0.001	52484	53055
<b>Mice</b>	448	10	<0.001	52534	53104
<b>Oxygen</b>	436	10	<0.001	52546	53117
<b>Lasers</b>	429	10	<0.001	52553	53124
<b>Plants</b>	415	10	<0.001	52567	53137
<b>Mother's genes</b>	411	10	<0.001	52571	53142
<b>Radioactivity</b>	408	10	<0.001	52574	53144
<b>GM fruit</b>	408	10	<0.001	52574	53144
<b>Cloning</b>	403	10	<0.001	52579	53150
<b>Experiments</b>	402	10	<0.001	52580	53150

*\*H0: constrained model; H1: direct effect for each item;*

Table 2.8 shows the results of the likelihood ratio tests. For the items ‘mice’, ‘cloning’, ‘mother’s genes’, ‘radioactivity’, ‘GM fruit’, and ‘lasers’, the model allowing only a direct effect is rejected in favour of the model allowing both intercepts and slopes to vary across groups. On the other hand, for items ‘plants’, ‘experiments’, ‘electrons’, ‘oxygen’, the model only freeing intercepts is not rejected by the likelihood ratio test. Looking at the BIC statistic for the direct effect models and the interaction models, the models where only intercepts are freed fit the data better. Put differently, there is evidence that the relationship between these items and the categorical latent variable varies by age; however, this is not extended to the item being interpreted completely differently by people from distinct age groups.

**Table 2.7 Model fit statistics – interaction for each item**

<b>Model</b>	<b>LR</b>	<b>df</b>	<b>p-value</b>	<b>AIC</b>	<b>BIC</b>
<i>Interaction*</i>	507	20	<0.001	52495	53127
<b>Electrons</b>					
<b>Mice</b>	495	20	<0.001	52507	53139
<b>Lasers</b>	448	20	<0.001	52554	53186
<b>Oxygen</b>	441	20	<0.001	52561	53193
<b>Cloning</b>	437	20	<0.001	52565	53197
<b>Mother's genes</b>	435	20	<0.001	52567	53199
<b>Plants</b>	430	20	<0.001	52572	53204
<b>Radioactivity</b>	430	20	<0.001	52572	53204
<b>GM fruit</b>	428	20	<0.001	52575	53206
<b>Experiments</b>	415	20	<0.001	52586	53218

*\*H0: constrained model; H1: interaction for each item*

**Table 2.8 Likelihood ratio tests - direct effect versus interaction for each item**

<b>Model</b>	<b>LR</b>	<b>df</b>	<b>p-value</b>
<i>Direct effect versus interaction*</i>	47	10	<0.001
<b>Mice</b>			
<b>Cloning</b>	34	10	<0.001
<b>Mother's genes</b>	24	10	<0.05
<b>Radioactivity</b>	22	10	<0.05
<b>GM fruit</b>	20	10	<0.05
<b>Lasers</b>	19	10	<0.05
<b>Plants</b>	15	10	0.13
<b>Experiments</b>	14	10	0.19
<b>Electrons</b>	10	10	0.46
<b>Oxygen</b>	5	10	0.89

*\*H0: Direct effect, H1: Interaction;*

The non-invariance found seems to be related to people from different age groups having different probabilities of giving a correct response, which is a measurement issue. It does not, however, appear to extend to a potentially different substantive interpretation of the question altogether by individuals from the three age groups selected here. Nonetheless, the

non-equivalence found is significant seeing that the most severe form extends to six out of the ten items.

## **2.4. Sensitivity Analysis**

Now that measurement non-invariance in relation to age has been discovered, as well as a class of people who say they don't know rather than give outright incorrect answers, it is important to ascertain whether employing a latent variable framework, which accounts for this, when predicting an attitudinal outcome makes a substantive difference. In other words, would we draw different conclusions about a substantive issue than if we just used the summed scale approach?

### **2.4.1 An attitudinal measure and science knowledge**

To test this, I look at attitudes to genetics. It is one of the very few attitudinal questions that is asked of both young people and adults in the Wellcome Trust Monitor Surveys. Respondents are asked in both Wave I and II: 'How optimistic are you about the possibility of improved healthcare as a result of genetic research. Are you...?'. The response options are: 1. Very optimistic; 2. Somewhat optimistic; 3. Not too optimistic; 4. Not at all optimistic. This was recoded into a binary variable by collapsing the first two and then the last two negative categories into one representing each type of attitude (Very/somewhat optimistic=83%; Not too/not at all optimistic=17%). The reference category represents positive attitudes towards genetics. A dummy variable is used for parsimony. This analysis is not intended as a robust way to shed empirical light on public attitudes to genetics. It is merely a test of what substantive conclusions can be drawn from using a latent variable framework compared to using a summed score approach.

Science knowledge was operationalised in three different ways. First, a three-class model accounting for the age-related measurement non-invariance explained in the previous

sections was used. Three classes are used instead of four because the fourth is not of great substantive interest. It was only found for the older group, the probabilities of response fit in between classes 2 and 3, without making much of a difference, so the more parsimonious 3-class model was used for the purpose of this short sensitivity analysis. Second, a summed scale was created from recoding all ten knowledge items into correct and incorrect responses, as is the usual approach in PUS literature ( $M=6.49$ ;  $SD=2.18$ ). Higher scores translate into more knowledge. This was also then recoded as a categorical variable with three categories: low (0-3); medium (4-7); high (8-10). The conversion to a categorical variable was done in order to render it more comparable to the latent class variable. The categories were created bearing in mind the nature of the latent classes (so the range of probabilities of response explained in the previous sections). This is, of course, not a perfect representation of the latent class variable, and it can never be that. Nevertheless, the ranges chosen for the categories seem to be a plausible approximation for the goal of this sensitivity analysis.

#### 2.4.2 Covariates

In both analyses gender is included as a covariate, in addition to age. Gender is included as a dummy variable with men as the reference category (female=56%; male=44%). The same age categories are included in the analysis using the summed knowledge measure with over 50s as the reference category. Educational attainment was not included because it is measured differently for adults and young people. The latter have a proxy measure of parents' highest educational achievement. It is implausible that an extra covariate would significantly alter the comparability of results. In the latent class model used here the science knowledge items are allowed to vary across the three age groups, thus accounting for the measurement non-equivalence identified in the previous step.

### 2.4.3 Analytical Approach

Binary logistic regressions are estimated. The logistic regressions using the summed knowledge measure (both the continuous and categorical version) are estimated in R. Because entropy is  $<0.8$ , the latent class logistic regressions are carried out in MPlus in a 1-step estimation of the latent classes. All indicators are free to vary across age groups, gender is added as another covariate, and the distal outcome is attitudes to genetics. The 1-step approach is preferred over extracting class membership and using that in a subsequent analysis. This is in line with recommendations in the literature. Extracting class information (to be used in a subsequent analysis) using the modal assignment rule with an entropy value lower than 0.8 (or as a general more robust estimation technique) would lead to smaller and biased standard errors and narrow confidence intervals. This renders significance tests unstable (Clark and Muthén, 2009).

### 2.4.4 Results

Table 2.9 shows results from the logistic regression models. The table shows odds ratios from the latent class model as well as from the analysis using the summed score approach. In model (1), the reference category is ‘medium knowledge’ for both the LCA and summed score analyses. In other words, the ‘don’t know’ class is compared to the medium knowledge class. The low knowledge category from the recoded knowledge scale is also compared to the medium knowledge category. In model (2) the reference category is the high knowledge class for the LCA analysis, and the high knowledge category for the summed score analysis using knowledge as categorical. Thus, the ‘don’t know’ and medium knowledge classes are compared to the high knowledge class.

**Table 2.9 Logistic regression models predicting attitudes**

<b>Attitudes to Genetics (ref: positive)</b>			
	<b>Ref: medium</b>	<b>Ref: high</b>	<b>Continuous</b>
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>
<b>LCA:</b>			
Don't know class(C1)	1.751***	4.494***	
Medium knowledge(C2)		2.567***	
<b>SUMMED SCORE:</b>			
Low knowledge	2.051***	3.841***	
Medium knowledge		1.873***	
knowledge scale			0.801***

**Note:**\*p<0.1; \*\*p<0.05; \*\*\*p<0.01; N=3,124.

The low and medium knowledge categories are compared to the high knowledge category. Knowledge is used as a scale ranging from 0 to 10 in model (3).

Odds ratios for both types of parameterisation of the model are very similar indeed. Across the board, people with lower levels of knowledge have higher odds of having a negative attitude towards genetics. The same story is confirmed in the model using knowledge as a continuous scale where going up a point on the knowledge scale decreases the odds of saying they are not optimistic about genetic research. In short, the odds ratios for both model parameterisations using the latent classes and comparable categories representing low,

medium, high knowledge are very similar. The results using knowledge as a scale also leads to the same substantive conclusions.

## **2.5. Conclusion**

This study set out to meet two primary goals. First, to uncover measures of information, lack of information, and misinformation, using the battery of ten science knowledge items administered in the Wellcome Trust Monitor Surveys. This is done in order to epitomise people's knowledge in a more nuanced, meaningful and representative manner. Second, to test whether the ten knowledge items exhibit measurement invariance across three age groups: 14-18 year olds, 19-49 year olds, and people aged over 50, respectively. This is done so as to avoid potentially erroneous conclusions about their levels of knowledge about science, biology and genetics in particular. A subsidiary goal set was to provide some recommendations regarding this battery for future waves of the survey.

A multi-group latent class model was employed to answer the following three research questions posited: 1) How many classes emerge from the data? Do the latent classes appear to be ordered in terms of their conditional probabilities? If yes, how many underlying dimensions can be inferred from the latent classes derived from the data?; 2) What is the extent of any potential age-related bias? In other words, are the items invariant across the three age groups? Are older respondents more likely to give a 'don't know' response; 3) For future Wellcome Trust Monitor Surveys, what recommendations can be made for the design of this battery of questions?

In order to answer the first question, I inspected the latent structure that emerged from the data, as well as the classes' substantive interpretation given by the pattern of conditional probabilities. Analysis of the pooled sample revealed a common 4-class structure as fitting

the data best in terms of the BIC. Separate analyses in the three age groups reveal a slightly different picture of the latent structures emerging from this data. Younger and middle-age respondents were classified into three classes, showing a tendency of increasing knowledge in terms of the apparent ordering of the conditional response probabilities. The first class represents a lack of information class because of very high conditional probabilities of answering 'don't know' to all items. Given this information, on the surface it seems that it might be a mistake to group these answers with incorrect answers, as is the most common practice in PUS studies. Empirically, they seem to represent different entities. However, the second class appears to be fairly knowledgeable, and probabilities of answering incorrectly are not large at all, which means it cannot represent a misinformation class. The third and last class is the most knowledgeable class, where conditional probabilities of answering correctly are very high.

Older respondents can be classified into four classes. Three classes have the same interpretation as for the other two age groups, but a fourth class comes somewhere between the first and second classes. Thus, it appears that their underlying knowledge spans across more levels. We could also say that it goes up in more increments than for the other two groups. Practically, this does not make a substantive difference; empirically, LCA is flexible enough to allow multi-group models to be estimated with different numbers of classes for one or some of the groups, so I could proceed to investigate measurement invariance (Kankaras et al., 2009). I can posit that two sub-dimensions of an underlying knowledge continuum emerge from the data inferred from how the classes appear to be ordered in terms of the conditional response probabilities. A 'don't know' class represents one, and the other dimension is represented by the other two classes (three for older respondents), which underpin a continuum of fairly good to high level of knowledge.



To answer the second question, a model where everything was held equal across groups was compared to one where everything was free to vary. Seeing as that fit the data better, I proceeded to search for the items that caused the most trouble (in terms of the amount of misfit provided). The results from the multi-group LCA show that all ten knowledge items are to some extent non-invariant across the three age groups. In other words, this battery of questions exhibits very strong age-related bias. This means that any observed differences might not indicate true differences. The models allowing the conditional response probabilities to vary across groups, regardless of the latent class, are the best fitting models. The likelihood ratio tests indicate this for ‘plants’, ‘experiments’, ‘electrons’, and ‘oxygen’. Thus, individuals have different probabilities of answering these items correctly based on what age group they belong to.

To answer the last part of the second question postulated, I examined the prior probabilities and the conditional probabilities in the ‘don’t know’ class. Older respondents in this class do not have much higher probabilities of saying they don’t know than younger respondents in the same class. However, 15% of older respondents are allocated to this class, compared to only 7% of young respondents. So it seems that many more older respondents have a propensity to give a ‘don’t know’ answer compared to younger respondents, thus confirming some of the observations made in political science literature (Ferber, 1966; Gergen and Bock, 1966; Glenn, 1969; Francis and Busch, 1975).

These items clearly need to be revised given the finding of a great amount of measurement non-invariance across the three age groups. There were also some differences in conditional probabilities for a few items in some classes. For instance, for ‘electrons’, the probabilities for younger and older respondents are reversed (high to give a correct response for young respondents, high to give a ‘don’t know’ answer for older respondents). Interestingly, the questions are administered separately to young respondents, as they are part of a boost

sample. Perhaps this is a reflection of how interviewers might treat them differently. Another possibility is that the information is just more fresh in their minds, given that most of them were probably still in school at the time of the interviews. Moreover, some of these items seem relatively independent across groups. This is based on the fact that the probabilities of giving a correct, incorrect, or ‘don’t know’ answer exhibit a high degree of similarity. In other words, the probability of giving a correct response is very close to that of giving an incorrect answer, or indeed of saying ‘don’t know’. The other items posing similar issues are ‘plants’, ‘mice’, ‘lasers’, ‘oxygen’, ‘mother’s genes’, and ‘experiments’. Coincidentally, except the last one, they are also the ones which are the largest source of error in terms of the BIC. Interestingly, only ‘lasers’ has ‘false’ as the correct answer. It has been shown that these items are more difficult for respondents and can sometimes be more cognitively burdensome, although they do help distinguish between respondents more effectively (see Stares, 2008). One practical recommendation would be to revise these items, and perhaps include more with ‘false’ as the response option as they seem to fare better here. Finally, another recommendation is to conduct more cognitive pretesting to understand how respondents engage cognitively with questions where ‘false’ is the correct response.

A sensitivity analysis was also conducted in order to ascertain whether the extent of measurement non-equivalence found in relation to age would bias substantive conclusions in analyses examining attitudes, for example. Binary logistic regressions were estimated using the latent class variable representing three incremental levels of knowledge, a summed knowledge scale, and a categorical variable representing low, medium, and high knowledge, akin to the latent class variable. Age and gender were used as covariates and measurement non-invariance was taken into account in the latent class logistic regression. The same substantive conclusions were reached using both types of modelling, so researchers can rest assured that using the summed score approach would not bias their conclusions regarding

issues of substantive interest. I reflect further on the implications of my findings in the ‘Overall Conclusion’, which is in Chapter 5.

In the next chapter I will examine education-based inequalities in science knowledge and biomedical self-efficacy. Science knowledge is used as a summed scale of correct and incorrect responses for two reasons. First, age-related differences in science knowledge are not the subject, so the measurement non-invariance in relation to age found here is not of concern; 2) the sensitivity analysis illustrated in Section 2.4 of the present chapter has shown that, substantively, either approach yields similar results. The latter gives me confidence that I only need to control for age in the subsequent analyses, and a latent variable framework is not necessary. Lastly, educational differences in science knowledge have been researched widely and have not been found to be an artefact of measurement.

## References

- Asparouhov, T., Muthén, B. (2012). Using Mplus TECH11 and TECH14 to test the number of latent classes. Mplus Web Notes: No. 14. May 22, 2012. Available online at: <http://www.statmodel.com/examples/webnotes/webnote14.pdf> [ Accessed January 20, 2013]
- Bartholomew, D., Knott, M., Moustaki, I. (2011). Latent variable models and factor analysis: a unified approach. Oxford: Wiley-Blackwell.
- Bauer, M. W., Durant, J., Evans, G. (1994). European Public Perceptions of Science. *International Journal of Public Opinion Research*, 6(2), 163-186.
- Bauer, M. W., Petkova, K., Boyadjieva, P. (2000). Public Knowledge of and Attitudes to Science: Alternative Measures That May End the "Science War". *Science, Technology, & Human Values*, 25(1, Winter), 30-51.
- Bollen, K. (1989). Structural equations with latent variables. New York: Wiley.
- Butt S, Clery E, Abeywardana, V., Phillips M (2009) Wellcome Trust Monitor 1: Survey Report. London: Wellcome Trust.
- Byrne, B. M., Shavelson, R. J., Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456-466.
- Byrne, B., Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34(2), 155-175.
- Clark S., Muthén B. (2009). Relating Latent Class Analysis Results to Variables not Included in the Analysis. Available online at: <http://www.statmodel.com/download/relatinglca.pdf> [Accessed July 12, 2017]

Clemence, M., Gilby, N., Shah, J., Swiecicka, J., Warren, D., Smith, P., Johnson, J., Hoolahan, F., D'Souza, J. (2013). Wellcome Trust Monitor Wave 2: Tracking public views on science, biomedical research and science education. London: Wellcome Trust.

Clogg, C. C., Goodman, L.A. (1984) Latent structure analysis of a set of multidimensional contingency tables. *Journal of the American Statistical Association*, 79, 762-771.

Converse, P.E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and Discontent*. New York: Free Press.

Dancey, L., Sheagley, G. (2013), Heuristics Behaving Badly: Party Cues and Voter Knowledge. *American Journal of Political Science*, 57: 312–325.

Dragow, F., Levine, M., Williams, E. (1985). Appropriateness measurement with polychotmous item response models and standardised indices. *Journal of Mathematical and Statistical Psychology*, 38, 67-86.

Durant, J., Evans, G., Thomas, G. (1989). The public understanding of science. *Nature*, 340, 11-14.

Ferber, R. (1966). Item Nonresponse in a consumer survey. *Public Opinion Quarterly*, 30, 399-415.

Francis, J., Busch, L. (1975). What we don't know about don't know. *Public Opinion Quarterly*, 40, 207-18.

Gergen, K., Back, K.W. (1965). Communication in the interview and the disengaged respondent. *Public Opinion Quarterly*, 30, 385-98.

Glenn, N. (1966). Aging, disengagement and opinionation. *Public Opinion Quarterly*, 33, 17-33.

- Glenn, N. (1969). White Gains from Negro Subordination. *Social Problems*, 14, 159-168.
- Hagenaars, J.A., McCutcheon, A.L. (2002). *Applied Latent Class Analysis*. Cambridge: Cambridge University Press.
- Kankaraš, M., Moors, G. (2009). Measurement equivalence in solidarity attitudes in Europe: Insights from a multiple-group latent-class factor approach. *International Sociology* 24, 557–579.
- Kuha, J. (2013). Multigroup latent variable modelling with the Mplus software (V6). Retrieved from [http://blogs.lse.ac.uk/lcat/files/2015/08/LCAT\\_Mplus\\_and\\_R.pdf](http://blogs.lse.ac.uk/lcat/files/2015/08/LCAT_Mplus_and_R.pdf) [Accessed January 24, 2013]
- Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., Rich, R. F. (2001). Misinformation and the currency of democratic citizenship. *Journal of Politics*, 62, 790–816.
- Lazarsfeld, P. F., Henry, N. W. (1968). *Latent Structure Analysis*. Boston: Houghton Mifflin.
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106–131.
- Luskin, R.C., Bullock, J.G. (2011). 'Don't Know' Means 'Don't Know': DK Responses and the Public's Level of Political Knowledge, *Journal of Politics*, 73(2), 547–557.
- McCutcheon, A. L. (1987). *Latent class analysis*. Newbury Park, CA: Sage Publications
- Miller, J. (1983). Scientific Literacy: a conceptual and empirical review. *Daedalus*, Spring , 29-48.

- Miller, J. (1992). Toward a scientific understanding of the public understanding of science and technology. *Public Understanding of Science*, 1 , 23-26.
- Miller, J. (1998). The measurement of civic scientific literacy. *Public Understanding of Science*, 7, 203-223.
- Miller, J. (2004). Public understanding of, and attitudes toward, scientific research: what we know and what we need to know. *Public Understanding of Science*, 13, 273-294.
- Muthén, L. K., Muthén, B. O. (1998-2010). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén.
- Nylund, K.L, Asparouhov, T., Muthén , B. (2007). Deciding on the number of classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling*, 14(4), 535-569.
- Sicinski,A.(1970). ‘Don’t Know’ Answers in Cross-National Surveys. *Public Opinion Quarterly*, 34, 126-129.
- Skrondal, A., Rabe-Hesketh, S. (2004). *Generalized latent variable modelling: multilevel, longitudinal and structural equation models*. Boca Raton, Florida: Chapman & Hall/CRC.
- Stares, S. R. (2008). *Latent trait and latent class models in survey analysis: case studies in public perceptions of biotechnology*. London: LSE.
- Sturgis, P., Allum, N. (2004). Science in society: re-evaluating the deficit model of public attitudes. *Public Understanding of Science*, 13, 55-74.

### **3. Educational Disparities, Biomedical Efficacy and Science Knowledge Gaps: can the Internet help us reduce these inequalities?**

#### **3.1. Introduction**

The cultural authority of science stems from the fact that modern institutions rely on expert knowledge to assess risks, make decisions, and establish legitimacy (Shapin, 2008; Liu and Priest, 2009). That being said, a society that is exhaustively reliant on expert knowledge without public involvement does not meet the democratic ideal. The basic tenet of a democracy is an active and informed citizenry, one which is equipped to comprehend basic scientific concepts and methods, and is aware of the main actors involved in science. This would enable those who are both willing and confident in their abilities to participate in the decision-making that has the potential to change their lives. They could, for instance, take part in public consultations regarding important policy matters, such as the regulation of medical research.

However, social and economic differences lead to unequal access to, and use of, information. This has been a pressing issue for years, with the World Economic Forum (Bilbao-Osorio et al., 2011) declaring that it should be recognised as one of the primary challenges that democratic governments have to tackle. The issue has been examined quite extensively using the knowledge gap theory method. In principle, the knowledge gap hypothesis, proposed by Tichenor, Donohue, and Olien (1970), tests whether an augmentation in media attention can increase knowledge disparities between social groups with different educational



backgrounds. The three authors posit that people would acquire knowledge at different rates, with more educated citizens becoming more informed at a faster rate than less educated members of society.

The existence of knowledge gaps has been tested in a variety of science domains, albeit less so in the biomedical/biotechnological realms. The measurement of knowledge has posed some particular issues. In essence, there is dissent in the field of Public Understanding of Science regarding how to assess understanding of scientific principles, facts, and actors. Factual knowledge measured with a series of true/false questions about textbook science knowledge has been the most commonly employed measure (e.g., Bonfadelli, 2002; Cacciatore, Scheufele and Corley, 2013; Ho, 2012). Other studies have shown preference for a measure of self-reported knowledge or perceived familiarity, where respondents are asked to rate how well informed they are about certain scientific issues (Kahan, Braman, Slovic, Gastil, and Cohen, 2009). And yet other researchers use them interchangeably (see Kahan et al., 2012). However, critics argue that the two concepts measure intrinsically different things and should not be employed together or as a replacement of each other. High factual knowledge does not always translate into high confidence in one's own abilities /information (Ladwig et al., 2012; Su et al., 2014; Anderson et al., 2014). Essentially, knowledge gaps may differ significantly depending on the type of measure we employ, which would have an impact on the inferences researchers make about different sections of society. This could then lead to different, potentially inaccurate, policies (Su et al., 2014).

We use a factual measure of knowledge of science, with a focus on biology and genetics. We also propose a more comprehensive measure than self-reported familiarity. The measure we put forth builds on two closely related concepts borrowed from political science literature, namely internal political efficacy and political information efficacy. By definition, internal political efficacy refers to personal beliefs regarding one's own ability to understand and

participate effectively in politics (Craig, Niemi, Silver, 1990; Morrell, 2003; Niemi, Craig, Mattei, 1991). Thus, we retain the element of self-reported understanding of biomedicine and we add a component of willingness to participate in the decision-making process regarding medical research. We argue that the new construct, which we call biomedical self-efficacy, or internal biomedical efficacy, offers a richer and better-rounded insight into people's engagement with biomedicine. The next section will provide more insight into the political science origins of the concept. To the authors' knowledge, biomedical self-efficacy has never been studied before. We found one PUS paper looking at nanotechnology efficacy, where it is treated as a predictor of risk perceptions, separate from perceived familiarity. However, a related notion, scientific citizenship, has been proposed in PUS (Mejlgaard 2009; Mejlgaard and Stares, 2010, 2013). Scientific citizenship differs from biomedical self-efficacy in that it postulates interdependencies between factual knowledge and participation (Mejlgaard, 2009). Efficacy, however, focuses on subjective assessments of knowledge and willingness to participate in debates. One caveat that needs to be made early on, however, is that preferences for participation have been found to not translate into actual participatory behaviour, as seen in Mejlgaard and Stares (2013). Ergo, any conclusions regarding an improvement in efficacy as meaning a better democracy where more people would actually participate are sadly not possible. We might be able to say that it can at least open up a more real possibility of them doing so.

This study aims to examine the knowledge gap hypothesis using factual knowledge and biomedical self-efficacy. One of the primary goals is to examine in detail the processes underpinning education-related factual knowledge and self-efficacy gaps. This study also investigates the role played by general Internet use and attention to medical stories on the Internet in narrowing or widening factual knowledge and biomedical self-efficacy gaps. There have been mixed findings in relation to the Internet's role in reducing knowledge

gaps. Some have found that the Internet can decrease (nanotechnology) knowledge gaps (Cacciatore et al., 2013), while others have advocated that the Internet enlarges gaps between low and high education groups (for example, Bonfadelli, 2002; Kim, 2008; Miller, 2001; partial evidence in Su et al., 2014). Some of these studies, however, only examine the impact that general Internet use has on knowledge gaps. This study adopts a more comprehensive approach to also examine how attention to medical stories online impacts education-based knowledge and internal biomedical efficacy gaps. An overview of the concept of internal political efficacy, on which we base that of biomedical self-efficacy, is imperative before any further exposition of the relationship between the Internet, biomedical self-efficacy, and science knowledge.

### 3.1.1. Biomedical Self-Efficacy

While biomedical self-efficacy is a measure first operationalised in this study, its conceptualisation is rooted in political science literature. Originally, the concept of "sense of political efficacy" was a unidimensional construct, defined as "the feeling that individual political action does have, can have, an impact upon the political process" (Campbell, Gurin and Miller, 1954:187). Subsequent theoretical and empirical work has demonstrated that the survey questions normally employed in measuring political efficacy actually represent two separate components. The first, internal efficacy, refers to individual beliefs regarding one's own competence to comprehend and take part effectively in politics. The second, external efficacy, measures beliefs about how responsive governmental bodies are to citizen demands (Craig, Niemi, Silver, 1990; Morrell, 2003; Niemi, Craig, Mattei, 1991). Kaid, McKinney, Tedesco (2007) propose a concept related to internal political efficacy, namely political information efficacy, which, unlike internal efficacy, has only one dimension. Information efficacy can be defined simply as citizens' confidence in their political knowledge and "its

sufficiency to engage in the political process" (Kaid, McKinney, Tedesco, 2007:1096; Tedesco, 2011).

The closest concept to biomedical efficacy, as conceptualised here, in PUS is scientific citizenship, which relates objectively measured knowledge to participatory behaviours (Mejlgaard, 2009; Mejlgaard and Stares, 2010, 2013). For instance, in a cross-national analysis using the European Survey on Europeans, Science and Technology, Mejlgaard and Stares (2010) do not find a definitive model that fully captures scientific citizenship across a multitude of European contexts. Their participation component includes questions about whether respondents discuss science, whether they read articles on science, and whether they attend public meetings/debates about science. The last question is more akin to the participation component of efficacy, as conceptualised in the current Chapter. The other two elements of the participation component of scientific citizenship are similar to those I use in Chapter 3, where I reconceptualise efficacy slightly (due to data constraints). They also find evidence that subjective competence is associated with that they call 'vertical participation' (which aims to influence policy agendas). Unlike scientific citizenship, efficacy focuses on a self-assessed and reported competence measure. In addition, its participation component can be thought of as aiming to influence policy seeing that it asks about willingness to take part in public consultations regarding the regulation of medical research. Therefore, it is not surprising to see that these two types of components appear to work together better than factual knowledge and participation. However, this is only a national sample, and there are no issues with cross-cultural comparability, which is likely also part of the reason.

Biomedical self-efficacy can thus be said to resemble the two efficacy measures used in political science more than scientific citizenship. It combines elements of internal efficacy and information efficacy. Biomedical self-efficacy has two components that work well in tandem. The first dimension or component refers to people's self-reported understanding of

three key terms in biomedicine, namely ‘DNA’, ‘Genetically modified’, and ‘Human Genome’, respectively. The second component relates to citizens' willingness to engage in public consultations regarding medical research. The ‘Method’ section details: a) the measurement properties and the statistical tests conducted to ascertain the empirical distinctness of biomedical self-efficacy; b) potential conceptual benefits, compared to the unidimensional self-reported familiarity measure.

To the authors' knowledge, at the time of writing, biomedical self-efficacy has not been measured before. Efficacy has, however, been measured in relation to nanotechnology in one PUS study. Anderson et al. (2014) study the online incivility and risk perceptions of emerging technologies, treating nanotechnology as a predictor of perceptions. They measure nanotechnology efficacy with two ten-point items: ‘Nanotechnology seems so complicated that a person like me can’t really understand it’, and ‘I would need more information about nanotechnology before I could make any decisions about it’. They also include a measure of self-reported familiarity, treating it as a separate construct altogether. The strength of this study stems from combining what we consider to be two elements of the same underlying concept. We argue that this creates a better-rounded measure of people’s self-reported cognitive engagement with medicine and biomedicine, as well as their disposition towards taking part in decision-making about medical research.

Furthermore, this study applies the knowledge gap hypothesis to both factual knowledge of science, with a focus on biology and genetics, and to biomedical self-efficacy. By comparing the impact of Internet use on both constructs, we expect to paint a more nuanced picture of the empirical realities of education-based inequalities. For instance, the discovery of efficacy gaps would have quite different policy implications. Knowledge does not always translate into confidence in one's abilities, or, indeed, a willingness to take part in the debate about medical research. Therefore, policy programmes only aiming to inform people cannot be the

sole answer to reducing inequalities. Indeed, the present study offers a clearer picture of the role of the Internet in both informing citizens and stimulating their confidence in their competence and willingness to participate in the civic debates about medical research. To put it differently, increased information, confidence, and participation should all be encouraged, for what is the first without the other when it comes to fulfilling our democratic ideals.

### 3.1.2. The Internet, Biomedical Self-Efficacy, and Science Knowledge

In the beginning, the knowledge gap hypothesis was only applied to newspaper use. Later, studies were published that examined gaps in television use. Seeing that the Internet occupies such a primordial role in today's information society, in the last decade or so, more and more studies have investigated its impact on inequalities.

The invention of the Internet has brought about changes in the media landscape that are consequential to the dynamic between audiences and science (Batts, Anthis, and Smith, 2008). In Western countries, more and more people access the news online. The particularities of the online environment have a different impact on public response to scientific information and new discoveries or technologies. These specificities are underpinned by information being conveyed in a new manner, and in new formats that offer innovative conduits of communication. Characteristics of the new information environment include, but are not limited to, hyper textuality (using links to connect to other web pages), interactivity, and multi-mediality (videos, images, etc.) (Brossard, 2013).

Of more importance to science audiences, however, is the fact that online information sources don't distinguish between 'news' and 'opinion'. For example, user-generated content (blogs, forums) is often an integral part of a news website. Thus, readers might find it difficult to disentangle the content in terms of what is scientific fact regarding, for instance, stem cell research, and what is the mere opinion of the blogger/forum participant.

Additionally, crowd-sourced content (such as Wikipedia) is also heavily present on the Internet. Coverage of scientific issues will differ significantly to that in more traditional mass media, such as newspapers. Another important feature of the Internet is the tailored nature of its content. In other words, people can choose to amend the terms of their online search for information to fit their pre-existing interests and knowledge. This might have interesting effects on citizens' self-efficacy and understanding of science. Importantly, according to the Office for National Statistics (2015, 2016) there is evidence showing that a majority of people living in the United Kingdom use the Internet. Currently, about 87.9% of adults in the UK use the Internet (ONS, 2016).

Back in 2012, when the data used in this study were collected, 68% of those aged over 16 in the UK used the Internet (ONS, 2012). Biomedicine and medical research are extremely prevalent in the online environment. Masic and Milinovic (2012) published a review of the multitude of online databases containing information about new and old discoveries in biomedicine. Their article shows that any person who wanted to access information about biomedicine and medical research online would have no difficulties in finding it. Moreover, Haustein et al. (2014) studied social media (Twitter) coverage of articles published on PubMed and Web of Science. They found evidence of very widespread coverage and circulation of biomedicine on Twitter, one of the most commonly used social media outlets. As a result, even if someone isn't actively looking for information related to biomedicine, they are still likely to encounter it at some point in time.

Be that as it may, the impact of the Internet on societal information discrepancies is still poorly understood and the evidence is mixed. This study aims to shed light on this important matter and further inform debates regarding the role of the Internet in reducing the deficit of information between the haves and the have-nots. Some researchers have found the Internet to be an equaliser in terms of information opportunities (i.e. Cacciatore et al., 2013; partial

evidence in Su et al., 2014). Yet others have argued that the Internet would in fact augment the digital divide, resulting in even larger information gaps between the rich and the poor. This is, presumably, because highly educated people would be better equipped to understand the information (see for example, Bonfadelli, 2002; Kim, 2008; Miller, 2001; partial evidence in Su et al., 2014). Lee (2009) showed that the gap in health knowledge caused by different education levels increased at higher levels of Internet use. He found evidence of a stronger relationship between Internet use and health knowledge amongst those highly educated, compared to those low on the education scale. Kim (2008) found a similar pattern for cancer knowledge. Cacciatore et al. (2013) showed that Internet use and attention to science online served to close the nanotechnology knowledge gap between high and low socio-economic groups. In a study on the effect of the pattern of news consumption on science knowledge levels, Su et al. (2015) found that use of online news increased science knowledge, independent of educational level.

Most of the studies applying the gap hypothesis only focus on factual knowledge measures. Su et al. (2014), however, compare factual and perceived knowledge. Moreover, they examine the effect of socio-economic status (SES), as determined by a combined measure of education and income, on factual and perceived nanotechnology knowledge gaps, and the differential influence of the mass media. Interestingly, they found that socio-economic status was only a predictor of factual knowledge, not of perceived knowledge. The authors also studied the impact of science blog use on SES-related factual and perceived knowledge gaps. As science blog use increased from 'no use at all' to 'low levels of use', the factual knowledge gap between high and low-SES groups diminished. At higher levels of use, the pattern reversed.

To our knowledge, no studies have yet examined biomedical, or other science-related, self-efficacy gaps. As a result, we turned to the political science literature for evidence related to



the relationship between Internet use, internal political efficacy and, a related concept, political information efficacy. Overall, the impact of Internet use on political efficacy and political information efficacy seems to be positive. For instance, Moeller, de Vreese, Esser, Kunz (2014) found that, for first-time voters in the Netherlands, actively participating in the communication process of political information online had a strong impact on their internal political efficacy. A number of studies have found a stable association between levels of Internet use, online interactivity, and increased internal political efficacy, as well as amplified political information efficacy (Tedesco, 2007; Tedesco, 2011). Others, however, have found that being exposed to information online is not enough (Scheufele and Nisbet, 2002; Nisbet and Scheufele, 2004). For instance, Nisbet and Scheufele (2004) found that web exposure alone did not produce significant changes in efficacy. The authors showed that political information obtained online increased political efficacy only in combination with discussion.

The problem with a lot of these studies has been an insufficient sample size (Boulianne, 2009; Kenski and Stroud, 2006). Given that this study benefits from a very good sample size, we expect our results to be more robust. In order to bypass the issue of Internet exposure not being sufficient, we have included a measure for attention to medical research stories online, which taps into citizens' cognitive engagement with the issue.

Based upon the review of literature on the relationship between the Internet, education, knowledge, and efficacy exposed above, we formulate the following three hypotheses:

**H1:** Higher education leads to a higher sense of biomedical efficacy and increased factual knowledge, compared to lower or no education.

**H2:** Internet use and attention to medical research stories online will reduce the biomedical efficacy gap between lowly and highly educated individuals.

**H3:** Internet use and attention to medical research stories will narrow the science knowledge gap between lowly and highly educated individuals.

### 3.1.3. Testing the Knowledge Gap Hypothesis

Tichenor et al. (1970) put forward two methods of testing for knowledge gaps. The first methodology proposes to look at the relationship between knowledge and education over time. If we observe an increase in the strength of association between the two, as more mass media content about a subject is disseminated, then we can say we have observed the production of a knowledge gap. As our data are not longitudinal in nature, this method is inaccessible. The second technique proposed by Tichenor et al. (1970) uses cross-sectional data and compares the correlation between education and knowledge for different issues that have distinct media coverage. If the correlation between education and knowledge is higher for those issues covered more intensely, then a knowledge gap has formed.

There is also a third method of testing for education-based gaps that relies on cross-sectional data and focuses on the individual, as opposed to aggregate measures of media exposure (used by Eveland and Scheufele, 2000, amongst others). If the correlation between socio-economic variables and knowledge is stronger for respondents with high mass media use than for those with a low level of media use, then we can say we are observing a knowledge gap. There are at least two advantages to adopting this approach. Firstly, this study measures media use at the individual level. By doing so, actual news exposure can be gauged more accurately compared to using the macro approaches described above. Secondly, it allows us to include a wider range of potential explanatory factors, while holding media use constant (Eveland and Scheufele, 2000; Eveland, Marton, and Seo, 2004).

## 3.2. Method

The data used in this study come from Wave II of the Wellcome Trust Monitor Survey. The Wellcome Trust Monitor is a triennial survey of adults and young people in the UK, based on a representative probability sample. Its main purpose is to provide a more accurate and nuanced picture of interest in, knowledge of, and attitudes to medical research, with a focus on biomedicine. Wellcome Monitor II took place in 2012 and interviews were conducted with 1,396 adults (aged 18+) and 460 young people (14-18 years old; 154 part of a ‘core’ sample and 306 of the ‘boost’ sample). We only use the adult sample because the questions of interest were not administered to the young sample. The response rate for Monitor II was 52.6% for the adult sample (Clemence et al., 2013).

A series of nested regression models were fitted to the data. In order to test the gap hypothesis, a set of interaction terms were created between our Internet use and attention variables, and education.

*The biomedical self-efficacy* measure was created by summing responses to four questions. Three of those assess people’s self-rated understanding of key terms in biomedicine, namely ‘DNA’, ‘Human Genome’, and ‘Genetically Modified’. The fourth evaluates their participation, or rather willingness to participate (used interchangeably henceforth), in public consultations about medical research ( $M=6.9$ ,  $SD=2.94$ ). In an effort to build a more robust measure of internal efficacy, we also explored associations between these four variables and other variables that were suspected to be related to them. These measured the following concepts: 1) people’s self-rated understanding of ethical issues in biomedical research; 2) whether responsibility regarding what to do about obesity and getting vaccinated lies with the public; 3) if anyone involved in decision-making about medical research should be well informed about the science involved; 4) willingness to participate in medical research.

Unfortunately, these avenues proved to be unfruitful as these variables did not exhibit any association with the four core measures of the construct.

Based on the moderate correlations found among the final four variables, an Exploratory Factor Analysis (EFA) was conducted. The result was one underlying dimension (one factor). Here, one dimension is used to denote the fact that the two components of efficacy, as theorised in the previous sections, represent one underlying construct empirically. Theoretically, however, as explained in the previous sections, efficacy was postulated to have two dimensions that work in tandem, but in the sense of components or elements: 1) self-reported understanding of key terms in biomedicine; 2) willingness to participate in public consultations. Empirically, these two components represent one underlying concept.

A Confirmatory Factor Analysis (CFA) was also carried out on the whole sample, and the hypothesized one-dimension structure (based on the initial exploratory analysis) fit the data very well (CFI=0.99, TLI=0.98, RMSEA=0.05). The diagram showing results from this analysis can be found in the Appendix to this Chapter. For the purpose of the regression models presented here, a scale was constructed by adding up responses to all four variables, which presented very good internal consistency (M=6.9; SD=2.97; Cronbach's alpha=0.72). Seeing that biomedical efficacy is a new construct, this study tested for measurement invariance across three educational groups: 'no qualifications' (28%), 'CSE, GCSE, A-level or equivalent' (15%), and 'higher education or equivalent' (57%), respectively. All efficacy items evidenced measurement equivalence, meaning that the items appear to work the same, empirically, and measure the same underlying concept, across the three educational groups.

However, measurement invariance testing is not the focus of this Chapter. It was only carried out to ensure that any educational differences in efficacy that may be observed are not an artefact of measurement, given that the operationalisation of this concept is new. Seeing that

it is merely a robustness check, extensive results are not presented here. Please see Table 3.3 in the Appendix to this chapter for model fit statistics for each type of invariance tested. These tests are similar to those conducted in the previous chapter, but they use a CFA framework. A description of all steps taken in the multi-group analyses can also be found in the Appendix.

As explained above, the first three questions ask respondents to self-assess their comprehension of three key terms in biomedical research. All three variables were ordinal, ranging from 'have not heard the term' (1) to 'very good' (5). The first probes respondents on their knowledge of the term 'DNA' (the sample median was 3 for 'some understanding'). The second queries them on their comprehension of the term 'Genetically modified' (the median was 1 for 'have not heard the term'). Lastly, they were asked to rate their understanding of the term 'Human Genome' (the median was 3 for "some understanding"). The third question asked them to choose from the following three statements: 'I don't want to be involved personally in public consultations on decisions about medical research'; 'I would like to become actively involved in public consultations about medical research'; 'I am already actively involved in public consultations on decisions about medical research'. It was treated as an ordinal variable with three categories, and recoded so that it followed an ascending order of involvement/willingness to be involved.

In PUS, a similar measure, namely perceived knowledge, has been used in comparable studies examining socio-economic related gaps. In light of this, two sets of correlations were compared in order to assess criterion validity. Firstly, correlations were examined between a measure using only the self-reported knowledge items and the key concepts internal efficacy should be related to. The second set of correlations inspected looked at how the same key terms were related to the new measure of internal efficacy, and whether these patterns of association were stronger than those for just the perceived knowledge measure.

The key terms identified in political science literature are: different forms of participation (different to the component of efficacy; these variables were found not to be a part of the efficacy construct in an EFA, as explained above), factual knowledge, and interest (Clarke, Kornberg, and Scotto, 2008). Even though the differences between the two sets of associations are very small (0.03 or 0.02; please see table 3.4 in the Appendix to this chapter), this result is consistent across all key concepts. Therefore, it reinforces the idea that the concept of internal efficacy, and its new element of desire to participate in decision-making in relation to regulating medical research, might potentially offer additional substantive insight into people's engagement with biomedical research.

*Factual science knowledge* was measured using ten questions. These are the same items used to measure science knowledge in Chapter 2. Nine of the questions are closed-ended items asking respondents if a series of statements about general science, biology and genetics, are definitely true, probably true, probably false, definitely false, or they can say they don't know. The nine statements were worded as follows:

- 1) Electrons are smaller than atoms. (T)
- 2) All radioactivity is man-made. (F)
- 3) All plants and animals have DNA. (T)
- 4) More than half of human genes are identical to those of mice. (T)
- 5) The cloning of living things produces identical copies. (T)
- 6) The oxygen we breathe comes from plants. (T)
- 7) Lasers work by focusing sound waves. (F)
- 8) By eating a genetically-modified fruit, a person's genes could also become modified. (F)
- 9) It is the mother's genes that determine the sex of the child. (F)

The tenth question tests the individuals' comprehension of controlled experimentation. They are asked: 'Suppose a drug used to treat high blood pressure is suspected of having no effect. On this card, there are 3 different ways scientists might use to investigate the problem. Which one do you think scientists would be likely to use?' Respondents are then asked to select one of the following options: 1) 'Talk to those patients that have used the drug to get their opinion?' 2) 'Use their knowledge of medicine to decide how good the drug is?' 3) 'Give the drug to some patients, but not to others, and then compare the results for each group?' 4) 'Don't know'.

Responses to all ten items are then recoded into dummy variables with '1' signifying correct answers, and '0' for incorrect answers. 'Don't know' answers are coupled with incorrect answers. This is informed by the sensitivity analysis conducted in Chapter 2. Responses are summed up to create a 0-10 scale, with high scores indicating a better understanding of science knowledge ( $M=6.4$ ,  $SD=2.17$ , Cronbach's Alpha= 0.63). Educational differences in science knowledge have been the subject of a plethora of studies since the 1970s, as discussed in the 'Introduction' and the subsequent review of literature, so there is no reason to suspect that it is a measurement invariance issue. Consequently, no such tests are performed for the science knowledge scale.

#### *Education and socio-demographic control measures*

Age, gender, education, and occupational status served as control measures in our analyses because studies have reported them as potential confounders of the relationship between Internet use, attention, and science knowledge (see for example Cacciatore et al., 2013). Age was measured as a continuous variable ( $M=51$ ,  $SD=17$ ). Gender was measured a dichotomous variable with male coded as '0' and female coded '1' (56% women; men are the reference category). Education was measured as an ordinal variable, ranging from 'no

qualifications' (coded '1') to 'higher education and above, or equivalent' (coded as '5'; the sample median value was '5'). Treating education in this manner is common practice in knowledge gap studies (see for example Cacciatore et al., 2013

Occupational status was measured as a three-category nominal variable, a recognised National Statistics Socio-economic Classification scale (NS-SEC3). The three categories were recoded into three dichotomous variables representing two occupational groups. The first dummy variable represents 'managerial and professional occupations' (38%). The second binary variable stands for 'intermediate occupations' (20%). According to the Office for National Statistics, the category of intermediate occupations from the NS-SEC 3 refers to clerical, sales, service, and intermediate technical occupations that do not involve planning or supervisory powers (ONS, 2010). The last category, 'routine and manual occupations' (41%), was left out, and represents the reference category. Two further controls related to the respondents' medical history were introduced. They are derived from a question asking why individuals had looked for information related to medical research online. Two dummy variables were created. The first was coded '1' for respondents who had looked for information because they'd suffered from a disease and '0' for those who hadn't looked for information, or had looked for it for different reasons, unrelated to illness, such as 'it was relevant to something I was studying', 'it was relevant to something I was doing at work', 'it is just an area that interests me'. The second dummy was coded the same, except '1' represents those who had looked for information because someone in their family had suffered from disease. Since the purpose of introducing them was solely to test the robustness of our results, and they didn't affect the results of interest (related to efficacy and knowledge gaps), we decided against retaining them in the final analyses. An ordinal measure of trust in information sources regarding medical research (combined measure of trust in Government, scientists working in universities, scientists working in industry,



medical charities) was also added initially. It was introduced as a control variable only, but it did not significantly alter the results and it is not of theoretical interest, so it too does not appear in the last version of the analysis. Please see Tables 3.6 and 3.7 in the Appendix to this Chapter for results including each of these additional controls.

### *Internet Use and Attention*

Researchers recommend using both attention and exposure measures (see for instance Eveland and Scheufele, 2000). Biomedical science and its applications can be quite controversial with the public (i.e. vaccinations, stem cell research, etc.). Haustein et al. (2014) have shown that scientific articles relating to biomedical discoveries had been widely discussed on Twitter. The effect of that wide circulation on one of the most popular social media sites could potentially help reduce disparities in knowledge and efficacy.

This study employs two measures of Internet use. The first is a dichotomous variable coded '0' for 'other' and '1' for 'Internet as the main source of news' (16% of respondents rely primarily on the Internet for information on news and current affairs). We used responses to two questions to derive this variable. The first asks respondents to say what sources of news provide them with the most information about public affairs, giving them the option to select multiple sources. The second asks what is their main source of news. However, if they only select one source for the first question, then they are not asked the subsequent question about which one is their main source of news. Those respondents are treated as missing instead. Initially, we created a dummy variable for those missing responses to the question inquiring into individuals' main source of news. We then cross-tabulated this against the first variable that asks them to identify one or several sources of news in order to identify the pattern of missing responses. This is how it emerged that those who were missing for the variable representing individuals' main source of news were those who had only chosen one option at

the first question. We then added responses identifying the Internet as the only source of news from the first question with responses from the second question. We subsequently dichotomised the result. We also measured weekly frequency of Internet use; this was an ordinal variable ranging from '1' 'Don't access the Internet' to '9' '13 hours or more' (the sample median was '2', representing 'less than an hour'). Given the large number of categories, this is used as a scale in all OLS regressions.

Attention to medical stories online is measured by asking respondents: 'Please think of the last time, before this interview, that you heard, saw or read something about medical research that you just happened to come across and had not been trying to find. Can you remember what that was?' If they say they remember, then they are presented with a list of possible media where they could have come across this information. Social media, e-mail, websites, non-science blogs, and science blogs, were all combined and recoded into one dummy variable coded '1' for those who remembered coming across information on the Internet (8.2%), and '0' for those who either did not remember or they had come across it on another medium (91.8%). Table 3.1 illustrates the distribution of all measures used in this study. For ordinal variables where the median was provided above, proportions are shown.

We use R 3.2.0/RStudio 0.99.893 for our moderated ordinary least squares regression models with mean centering; we use Mplus 7.3 (Muthén & Muthén, 1998-2010) for the confirmatory factor and measurement invariance tests. In R, we use the 'pequod' package version 0.0-5 to fit our moderated regression, and to get the graphics for our two and three-way interaction terms (Mirisola and Seta, 2016). Variables were introduced in what appeared to be the logical order. No claims to causality are made seeing that we are only using cross-sectional data. Model 1 shows results for education, along with other socio-demographic

**Table 3.1 Descriptive statistics**

<b>Variable name</b>	<b>Mean/%</b>	<b>S.D.</b>
<b>Biomedical self-efficacy</b>	6.9	2.94
<b>Science knowledge</b>	6.4	2.17
<b>Age</b>	51	17
<b>Education</b>	5	
<b>Managerial occupations</b>	38	
<b>Intermediate occupations</b>	20	
<b>Routine occupations</b>	41	
<b>Gender (female)</b>	56	
<b>Web main</b>	16	
<b>Web frequency</b>	2	
<b>Attention (pays attention)</b>	8.2	

variables as controls. It makes sense to first establish the presence of education-based gaps, controlling for other SES factors, before testing the impact of Internet use and attention on those. Model 2 adds the Internet use and attention variables. Table 3.5 in the Appendix shows results from initial regressions testing the relationship between: 1) education and efficacy; 2) education and factual knowledge; 3) Internet use and attention, and efficacy; 4) Internet use and attention, and factual knowledge. Model 3 includes the three-way interaction terms between education, Internet use, and attention variables in order to test for the gap hypothesis. All continuous/ordinal variables, namely age, education, Web frequency, and science knowledge, have been centered at their mean. This makes the interaction effects independent of the main effects. Also, it changes the interpretation of the intercept, which now represents the change in efficacy or factual knowledge, when all other variables are held at their mean.  $R^2$  is used to assess how well each model fits our data.

### 3.3 Results

Table 3.2 shows results from our ordinary least squares regression models predicting, in turn, biomedical efficacy, and factual knowledge, respectively. Tables were created using package ‘texreg’ (Leifeld, 2013). Graphs were created using package ‘pequod’ (Mirisola and Seta, 2016). Standard errors appear in brackets.

Models 1 show results for our socio-demographic controls, predicting, in turn, biomedical efficacy and factual knowledge, respectively. For the efficacy model, 19% of the variance in efficacy can be explained by socio-demographic characteristics. For the factual knowledge model, 23% of the variance in objective competence is explained. There is a very small negative association between age and efficacy, controlling for everything in the model. There also seems to be a very small decrease in factual knowledge levels with age, and this result is statistically significant. However, given how small the association is, this result does not carry a lot of substantive significance. Holding everything else in Model 1 constant, women are less confident in their abilities to understand key terms in biomedicine and less willing to be actively involved in the decision-making regarding medical research. Moreover, women also appear to have a lower understanding of factual and experimental science than men.

As expected, we find evidence that supports **H1**, whereby individuals who are highly educated feel more efficacious and are more knowledgeable than those with Alevels/GCSEs/Olevels or no qualification, controlling for age, gender, and occupational status. This indicates presence of education-based inequalities/gaps. Unsurprisingly, a similar pattern can be observed for individuals in managerial occupations compared to those doing routine and manual jobs. The former feel more efficacious and are more knowledgeable too, everything else held equal.

Models 2 present results for the model introducing the Internet use variables. The amount of variance in self-confidence that is explained increases by 2 percentage points with the addition of these variables. It only increases by 1 percentage point for the factual measure. Results for the demographic controls remain consistent, with very slight variations, after controlling for the Web as the main source of news, frequency of Web use, and whether individuals remember coming across information related to medical research online (which we call ‘attention’). Notably, the coefficient for education drops very slightly once Internet use and attention is taken into account. Controlling for all demographics, those whose primary source of news is the Internet have a higher sense of efficacy than individuals relying on other forms of media. However, this result is not statistically significant. Moreover, the same applies to factual knowledge, but this is also not statistically significant. The more someone uses the Internet, the more efficacious they feel, holding everything else constant. This result is statistically significant. The same is true for levels of factual knowledge.

There is also evidence of an association between attention to medical stories online and biomedical efficacy. Controlling for all demographics and other Internet use variables, respondents who report that they remember coming across information related to medical research on the Web have heightened self-reported knowledge and desire to be involved in decision-making regarding medical research. Remembering that one has come across information on the Internet also has a positive impact on respondents’ levels of actual knowledge. As it transpires, using the Internet a great deal in general and paying some attention to the content has great benefits for people’s sense of efficacy regarding biomedicine.

Models 3 illustrate results from the regression models introducing the two and three-way interaction effects between educational level, frequency of Web use, and attention to medical

stories online, in order to test for the gap hypothesis. Even though results for the two sets of interactions are presented together (for simplicity), the coefficients for the two-way interactions come from a model that was run separately. The model with three-way interactions does control for two-way interactions.

This final model explains twenty-two percent of the variability in efficacy feelings, and twenty-four percentage of the volatility in objectively measured knowledge. The three-way interaction is only statistically significant for biomedical efficacy. However, the two-way interaction between education and Web frequency is statistically significant, and it offers insights into whether using the Internet more frequently has an impact on educational-based efficacy and knowledge gaps.

Figure 3.1 portrays the two-way interaction between education and frequency of Internet use on biomedical efficacy. Firstly, as expected, there is a stark difference in efficacy between the low and high education groups. At higher levels of Internet use, those with low or no education gain the most in confidence, while those more highly educated only experience a minute gain. Thus, the education-based efficacy gap seems to diminish gradually with increased use of the Internet. As discussed above, both frequency of use and attention are important in assessing the impact of mass media on other social phenomena.

Figure 3.2 portrays the three-way interaction between education, frequency of Internet use, and attention on biomedical efficacy. Firstly, as expected and as seen above, there is a clear difference in efficacy between the low and high education groups, independent of frequency of use or whether they pay attention or not. Moreover, amongst the lowly educated, it seems that paying attention to medical stories online, as well as increased general use, offer rewards in terms of their sense of biomedical efficacy. In other words, individuals in this group become are more efficacious to start with; they only get a small boost in confidence if they

use the Internet a great deal, provided a minimum cognitive engagement with the content. Those who do not pay attention at all start on a much lower level, but they experience a greater gain in confidence when their Internet use level augments. The gap between the lowly educated who pay attention and have higher confidence and those who do not pay attention and start with lower confidence closes significantly. This may mean that these individuals only need to be actively using the Internet more often to start feeling more confident in their abilities, and more willing to participate in the democratic process regarding medical research.

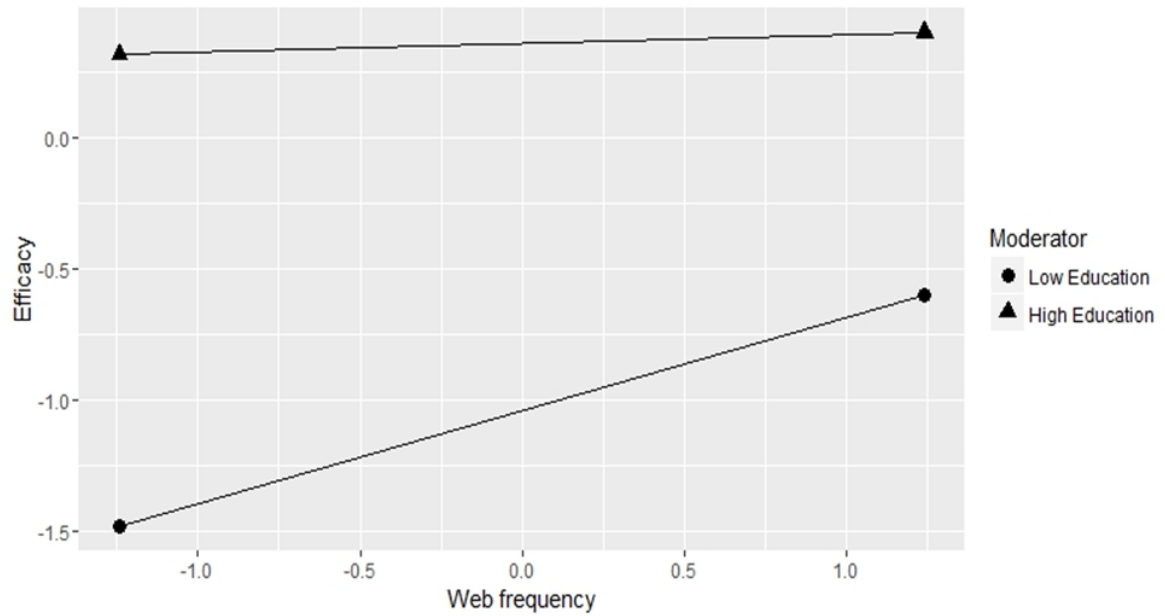
This configuration does not exactly replicate for the highly educated group. There is almost no difference between those with higher education who pay attention and those who do not. Respondents who do not pay attention to medical stories online experience a slight drop in efficacy at higher level of general Internet use. Highly educated individuals, who report having some semblance of a cognitive engagement with the content, adopt greater efficacy at increased levels of Internet use.

In line with **H2**, the difference between the highly educated group, who do not engage with the content, and the lowly educated group, who do engage, is slightly reduced at high levels of general Internet use, holding everything else constant. Additionally, at raised levels of Internet use there is a significant diminution of the gap between the highly and lowly educated, both of which have no recollection of seeing medical stories anywhere on the Internet. It seems that merely using the Internet a lot, in general, is enough to augment the confidence and the desire to participate of those low or no education.

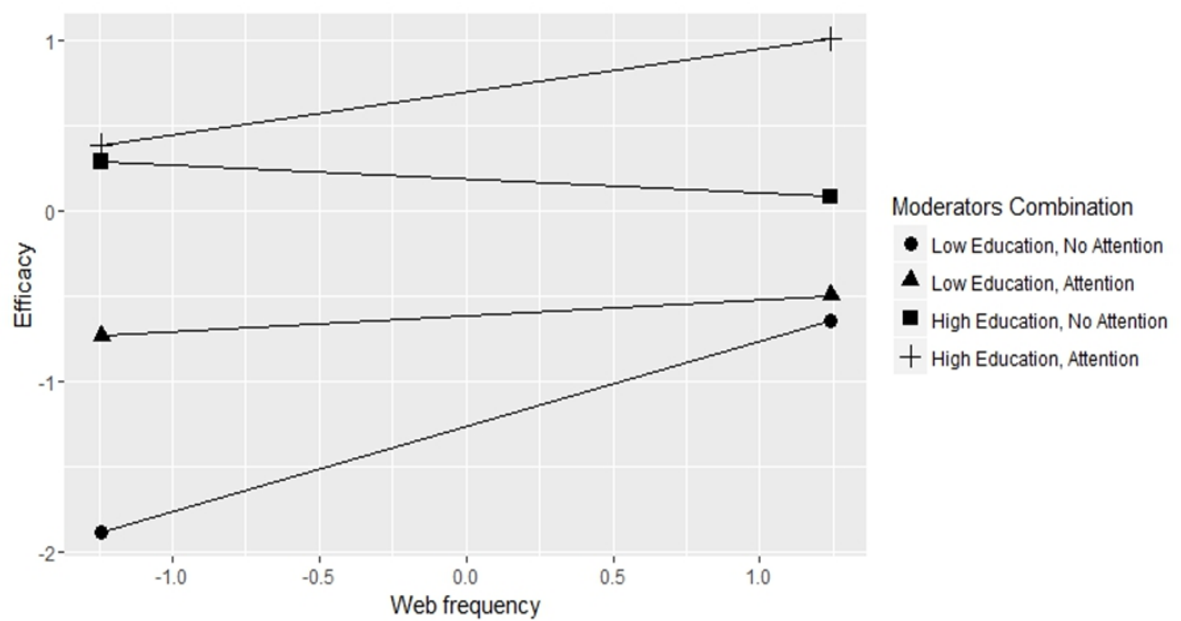
The gap between the lower and higher education groups, both of which pay attention to online content related to medical research, widens, which is more in line with the traditional knowledge gap hypothesis. Curiously, however, those who are lower on education and do

not pay attention to medical narratives online, catch up, albeit very slightly indeed, to those who are highly educated and pay attention, in tune with **H2**. However, those in the

**Figure 3.1 Two-way interaction : frequency of Web use and efficacy**

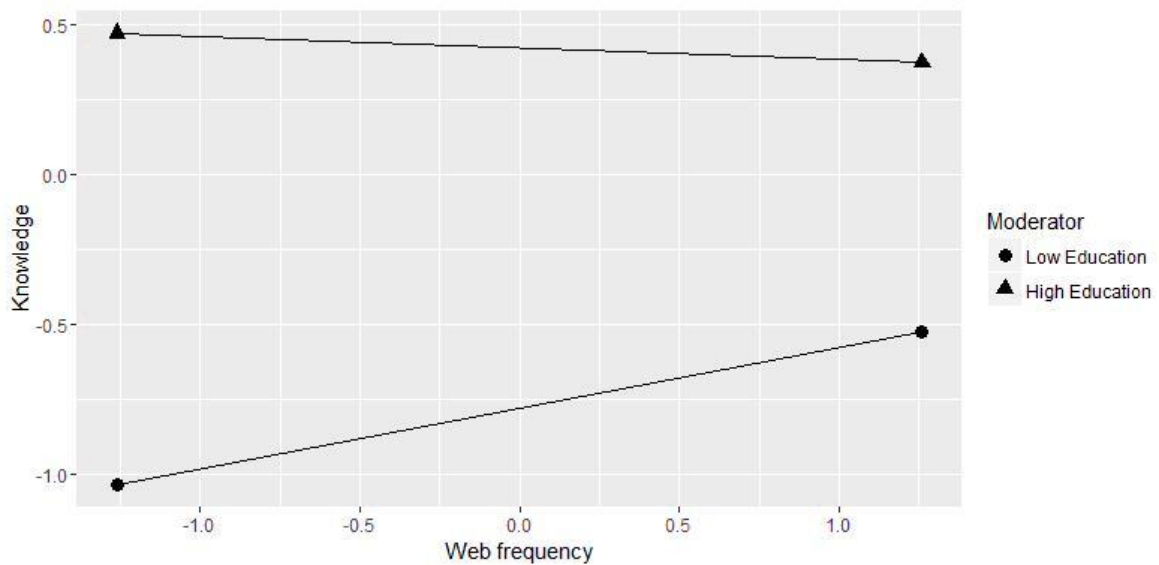


**Figure 3.2 Three-way interaction: frequency of Web use, attention, efficacy**





**Figure 3.3 Two-way interaction : frequency of Web use and knowledge**



more educated group who pay attention still have a higher sense of efficacy concerning biomedicine. In conclusion, we find different scenarios of some gaps closing and another widening slightly, so there is mixed support for **H2**.

Lastly, **H3** receives support, as exhibited in Figure 3.3. Controlling for everything else in the model, the knowledge disparity between those who are highly educated and those who are not well educated dwindles at high levels of general Internet use. Interestingly, those with higher education experience a minor drop in factual knowledge the more they use the Internet. This sits in contrast to what happened with efficacy levels (which rose slightly) in the two-way interaction between education and frequency of Internet use. Conversely, those with lower levels of education, experience a significant gain in understanding of science when Internet use increases. Overall, it seems that the lower education group gain the most from using the Internet.

**Table 3.2 OLS models predicting Biomedical Self-Efficacy and Science Knowledge**

	Biomedical self-efficacy			Science Knowledge		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	-0.27 (0.15)	-0.40 * (0.15)	-0.34 * (0.16)	-0.16 (0.11)	-0.22 (0.11)	-0.18 (0.11)
Demographics						
Age	-0.01 (0.00)	0.00 (0.01)	0.00 (0.01)	-0.01 *** (0.00)	-0.01 (0.00)	-0.01 (0.00)
Female	-0.60 *** (0.16)	-0.54 *** (0.16)	-0.55 *** (0.16)	-0.69 *** (0.11)	-0.67 *** (0.11)	-0.65 *** (0.11)
Education	0.46 *** (0.05)	0.43 *** (0.05)	0.40 *** (0.05)	0.36 *** (0.04)	0.35 *** (0.04)	0.34 *** (0.04)
Manager	1.44 *** (0.19)	1.32 *** (0.19)	1.30 *** (0.19)	1.10 *** (0.13)	1.05 *** (0.13)	1.03 *** (0.14)
Intermediate	0.22 (0.21)	0.21 (0.21)	0.17 (0.21)	0.64 *** (0.15)	0.63 *** (0.15)	0.60 *** (0.15)
Internet use						
Web main		0.29 (0.23)	0.33 (0.23)		0.15 (0.16)	0.16 (0.16)
Web freq.		0.23 ** (0.07)	0.19 * (0.08)		0.10 * (0.05)	0.08 (0.05)
Attention		1.15 *** (0.29)	1.05 ** (0.38)		0.44 * (0.21)	0.51 (0.28)
Two-way interactions						
Educ. X Web freq.			-0.04 * (0.04)			-0.06 * (0.03)
Educ. X Attention			-0.01 (0.28)			-0.27 (0.20)
Web freq. X Attention			-0.25 (0.26)			0.21 (0.19)
Three-way interactions						
Educ. X Web freq. X Attention			0.38 * (0.18)			0.07 (0.13)
$R^2\%$	19	21	22	23	24	24

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ ; N efficacy models=1166; N knowledge models = 1233; S.E. in brackets;

This reduction of the educational-based gap in knowledge at high levels of Internet use signifies that individuals in the lower education group do not just become more confident in their abilities; they also gain a slightly better understanding of science. It is an encouraging picture and further research avenues should explore a breakdown of online activities and how they relate to people's sense of efficacy as well as their knowledge of scientific facts and methods. Data limitations, due to the small number of observations for each type of activity respondents are asked about prevent us from pursuing that here.

### **3.4. Conclusion**

In this study we embarked on an attempt to understand the impact of Internet use and attention to medical stories online on education-based differences in factual science knowledge and biomedical self-efficacy. To the authors' knowledge at the time of writing, biomedical efficacy had not been measured before. A related concept of scientific citizenship, which combines factual knowledge and participation, was proposed by Mejlgaard (2009), and the Mejlgaard and Stares (2010, 2013). In their attempt to measure scientific citizenship, they also looked at perceived knowledge and participation. Their discovery that perceived knowledge is related to participation vis-à-vis policy matters informed the conceptualisation of efficacy. This was further informed by the political science literature on internal political efficacy and political information efficacy. Thus, our new measure was postulated to have two components: self-reported familiarity with key terms in biomedicine and willingness to participate in public consultations regarding the regulation of medical research. They were found to represent one construct, an underlying self-confidence with regards to biomedicine, which was termed biomedical self-efficacy. Perceived knowledge on its own has been widely used in PUS studies. Biomedical efficacy

contributes by taking into account people's wider sense of confidence, encompassing participatory attitudes.

Since the 1970s, when Tichenor et al. first proposed the Knowledge Gap Hypothesis, there has been a plethora of studies examining educational differences in knowledge and the mass media's effect on these. However, most of these studies have been conducted in the US. The current study has helped further our understanding of education-based differences in knowledge and efficacy in a UK context.

What is more, the democratic deficit in information has persisted, and new forms of information propagation have appeared. The Internet has revolutionised the way information is disseminated. Be that as it may, we still do not have a good grasp on how its use impacts upon science knowledge gaps in the UK, with the evidence on this being mixed at best. Lastly, there seems to be no evidence on how it affects self-confidence with regards to biomedicine. In order to assess Internet use in a more robust way, in line with recommendations in mass media effects literature, we also take into account reported cognitive engagement with online content regarding medical research (i.e. Eveland and Scheufele, 2000). Seeing that only 8.2% of our sample reported such an engagement, we think it is unlikely to have been severely skewed by a social desirability bias.

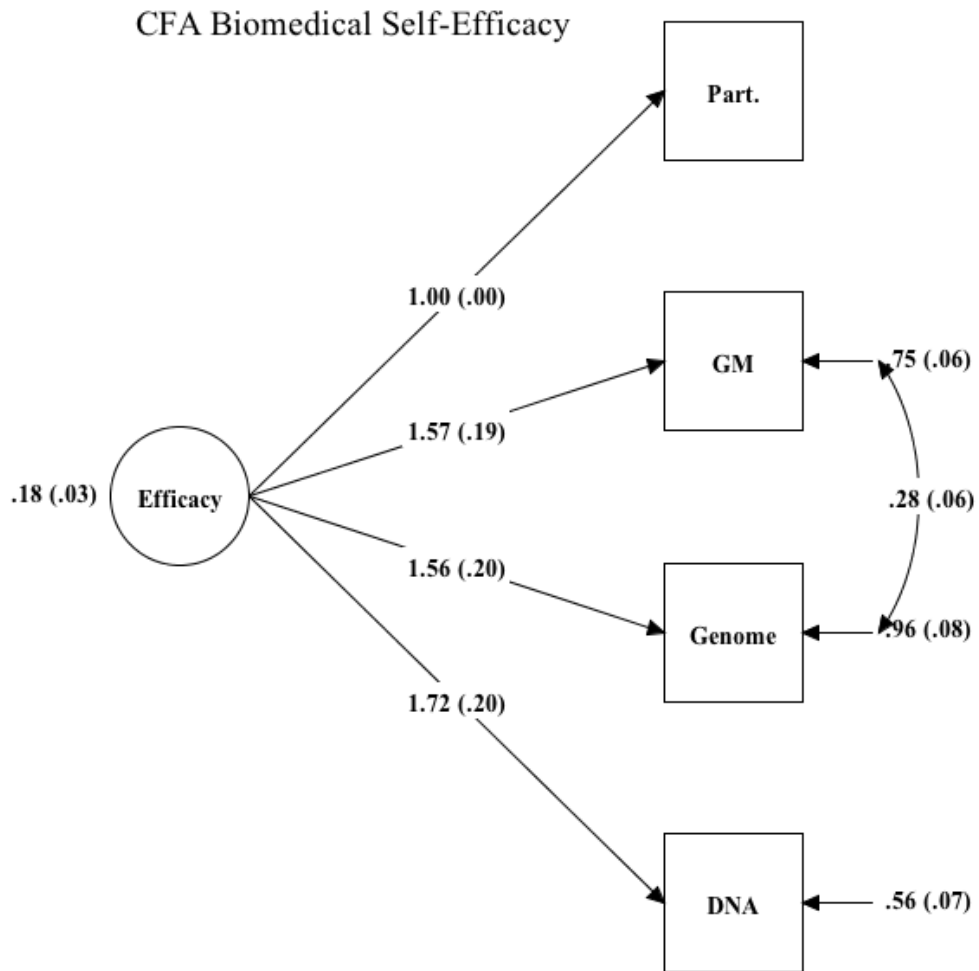
Moderated ordinary least squares regressions were employed to test the following three hypotheses: 1) Higher education leads to a higher sense of biomedical efficacy and increased factual knowledge, compared to lower or no education; 2) Internet use and attention to medical research stories online will reduce the biomedical efficacy gap between lowly and highly educated individuals; 3) Internet use and attention to medical research stories will narrow the science knowledge gap between lowly and highly educated individuals.

We find evidence that lends support to the first hypothesis. Those who are in the higher education group are found to be more efficacious as well as more knowledgeable. With regards to our second hypothesis, the results are mixed, with some gaps closing and others widening. In terms of their sense of biomedical efficacy, those who have lower or no education and say they pay attention to medical stories online catch up to those who are highly educated and do not pay attention, at high levels of Internet use. The most promising result indicates that merely using the Internet a great deal, without saying that they pay attention, leads to a significant diminution of the efficacy gap between those who are highly educated and those who are not. On the opposite side, when both education groups say they pay attention, the efficacy gap between them widens when they use the Internet a lot.

The third hypothesis received support, whereby the knowledge gap between lower and higher educated respondents is reduced when the former report using the Internet a great deal. In terms of broader socio-economic traits, women are found to be less confident and less knowledgeable than men. Individuals in managerial positions are much more confident and knowledgeable than people in routine and manual jobs. Interestingly, the associations between occupation and efficacy, as well as occupation and knowledge are stronger than those for education. This opens up a potential avenue for future research into whether mass media can help bridge those gaps.

Further reflections on the theoretical and empirical contributions of this Chapter can be found in the ‘Overall Conclusion’.

### 3.5. Appendix



## **Measurement Invariance of Biomedical Efficacy Across Three Educational Groups: a CFA Approach:**

The following steps were taken to test measurement equivalence of biomedical self-efficacy across three educational groups: ‘no qualifications’, ‘CSE, GCSE, A-level or equivalent’, and ‘higher education or equivalent’, respectively:

- 1) In order to test construct/structural invariance, a 1-factor model was fitted in each educational group separately; model fit statistics indicated that the hypothesized one-dimensional construct fits well in each sub-population. Thus, we can proceed to test for scalar invariance using multi-group models. CFI, TLI, RMSEA are used to assess model fit (Bartholomew, Knott, and Moustaki, 2011).

In Chapter 2, I initially started with a fully constrained multi-group LCA model, and then freed parameters in turn. A Weighted Least Squares (WLSMV) estimator was used; this is one of the recommended estimators in the MPlus User Guide when indicators are categorical. WLSMV requires nested models to be compared using the DIFFTEST procedure. This, in turn, requires that the first model be the least restrictive model (Muthén & Muthén, 1998-2010). Muthén & Muthén (1998-2010) also recommend that measurement invariance tests for a CFA with categorical indicators be conducted in two steps only. This brings me to step 2 in testing for invariance:

- 2) In order to test for scalar invariance, all thresholds and loadings are constrained to be equal across the three educational groups in a multi-group CFA model. I then compare this model to the less restrictive model using the DIFFTEST procedure, as recommended in the MPlus User Guide. DIFFTEST is a chi-square test assessing change in model fit when a more restrictive model is fitted; a significant p-value means that the restriction worsens model fit, so the less restrictive model should be

accepted instead (Muthén & Muthén, 1998-2010). Model fit does not worsen with the additional restriction (constraining thresholds and loadings to be equal across groups). Therefore, a model that is equivalent across groups seems more adequate.

**Table 3.3 Model fit statistics for measurement equivalence tests**

Model	<i>DIFFTEST</i> $X^2$	p-value	CFI	TLI	RMSEA
<i>Construct equiv.</i>					
<i>No qualifications</i>			0.99	0.99	0.09
CSE, GCSE, A-level			0.99	0.99	0.08
Higher education			0.99	0.99	0.08
<i>Fully unconstrained (baseline)</i>			0.99	0.99	0.05
<i>Fully constrained</i>	59.47	>0.05			



**Table 3.4 Assessing the Criterion Validity of Biomedical Efficacy**

		Self-efficacy	Self-reported knowledge
Self-Efficacy	Pearson Correlation	1	.987**
	Sig. (2-tailed)		.000
	N	1304	1304
Just self-reported knowledge	Pearson Correlation	.987**	1
	Sig. (2-tailed)	.000	
	N	1304	1360
Trust (information sources)	Pearson Correlation	.024	.030
	Sig. (2-tailed)	.398	.285
	N	1207	1257
Fictional book about science	Pearson Correlation	.329**	.317**
	Sig. (2-tailed)	.000	.000
	N	1300	1356
Factual book about science	Pearson Correlation	.403**	.384**
	Sig. (2-tailed)	.000	.000
	N	1301	1356
Science centre	Pearson Correlation	.281**	.272**
	Sig. (2-tailed)	.000	.000
	N	1303	1359
Participated in medical research	Pearson Correlation	.203**	.184**
	Sig. (2-tailed)	.000	.000
	N	1300	1356
Willing to participate in medical research again	Pearson Correlation	.279**	.250**
	Sig. (2-tailed)	.000	.000
	N	1265	1309
Medical research will improve lives	Pearson Correlation	.164**	.155**
	Sig. (2-tailed)	.000	.000
	N	1273	1324

\*\* Correlation is significant at the 0.01 level (2-tailed).

**Table 3.5 Education and Internet use on efficacy and knowledge**

	Efficacy		Knowledge	
	Model1	Model2	Model3	Model4
Intercept	0.00	-0.11	0.00	-0.08
Education	0.63 ***		0.50 ***	
web main		0.70 **		0.49 **
web freq.		0.41 ***		0.28 ***
Attention		1.63 ***		0.83***
<i>R</i> <sup>2</sup> %	14	7	16	5

Note:\*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05

**Table 3.6 Models predicting knowledge with additional controls**

	Factual Knowledge		
	(1)	(2)	(3)
Intercept	-0.218* (0.115)	-0.243** (0.116)	-0.185 (0.115)
Age	-0.002 (0.004)	-0.003 (0.004)	-0.001 (0.004)
Female	-0.643*** (0.113)	-0.616*** (0.112)	-0.606*** (0.112)
Education	0.320*** (0.038)	0.322*** (0.038)	0.331*** (0.038)
Manager	0.971*** (0.136)	0.999*** (0.136)	0.981*** (0.136)
Intermediate	0.517*** (0.153)	0.514*** (0.153)	0.517*** (0.154)
web main	0.172 (0.163)	0.146 (0.163)	0.175 (0.164)
web freq.	0.044 (0.054)	0.043 (0.054)	0.051 (0.054)
Attention	0.368 (0.267)	0.368 (0.266)	0.430 (0.267)
Fam.disease	0.360** (0.150)		
Resp.disease		0.504*** (0.165)	
Trust			0.016 (0.021)
Educ. X web freq.	-0.067** (0.028)	-0.067** (0.028)	-0.067** (0.028)
Educ. X Attention	-0.192 (0.195)	-0.212 (0.195)	-0.200 (0.196)
web freq. X Attention	0.181 (0.182)	0.155 (0.181)	0.163 (0.182)
Educ. X web freq. X Attention	0.069 (0.129)	0.089 (0.128)	0.071 (0.129)
<i>R</i> <sup>2</sup> %	23	23	22

Note:\*p<0.1; \*\*p<0.05; \*\*\*p<0.01; N=1233;

**Table 3.7 Models predicting efficacy with additional controls**

	Efficacy		
	(1)	(2)	(3)
Intercept	-0.469*** (0.161)	-0.451*** (0.163)	-0.361** (0.162)
Age	0.006 (0.005)	0.005 (0.005)	0.007 (0.005)
Female	-0.628*** (0.158)	-0.532*** (0.158)	-0.537*** (0.159)
Education	0.367*** (0.055)	0.387*** (0.055)	0.401*** (0.055)
Manager	1.230*** (0.189)	1.292*** (0.190)	1.290*** (0.192)
Intermediate	0.098 (0.215)	0.096 (0.217)	0.122 (0.218)
web main	0.387* (0.229)	0.350 (0.231)	0.389* (0.232)
web freq.	0.117 (0.077)	0.132* (0.078)	0.150* (0.078)
Attention	0.817** (0.359)	0.903** (0.361)	0.986*** (0.362)
Fam.disease	1.110*** (0.208)		
Resp.disease		0.770*** (0.231)	
Trust			-0.032 (0.031)
Educ. X web freq.	-0.095** (0.040)	-0.096** (0.040)	-0.099** (0.041)
Educ. X Attention	0.010 (0.267)	-0.041 (0.269)	-0.031 (0.270)
web freq. X Attention	0.071 (0.244)	-0.004 (0.245)	-0.012 (0.247)
Educ. X web freq. X Attention	0.353** (0.176)	0.398** (0.177)	0.391** (0.178)
<i>R</i> <sup>2</sup> %	23	22	21

Note:\*p<0.1; \*\*p<0.05; \*\*\*p<0.01; N=1166;

## References

- Anderson, A., Goel, S., Huber, G., Malhotra, N., & Watts, D.J. (2014). Political Ideology and Racial Preferences in Online Dating. *Sociological Science*, 1, 28-40.
- Baker, L., Wagner, T.H., Singer, S., Bundorf, M.K. (2003). Use of the Internet and E-mail for Health Care Information. *The Journal of the American Medical Association*, 289, 2400-2406.
- Batts, S.A., Anthis, N.J., Smith, T.C. (2008). Advancing science through conversations: Bridging the gap between blogs and the academy. *PLOS Biology* (p. 240). <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0060240>. [Accessed August 10, 2016]
- Bilbao-Osorio, B., Blanke, J., Browne, C., Crotti, R., Fidanza, B., Drzeniek Hanouz, M., et al. (2011). The Global Competitiveness Report: 2011-2012. In K. Schwab (Ed.). Geneva, Switzerland: World Economic Forum.
- Bonfadelli, H. (2002). The Internet and Knowledge Gaps. A Theoretical and Empirical Investigation. *European Journal of Communication*, 17, 65-84.
- Boulianne, S. (2009). Does Internet Use Affect Engagement? A Meta-Analysis of Research. *Political Communication*, 26, 193-211.
- Brossard, D. (2013). New media landscapes and the science information consumer. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 14096-14101.
- Cacciatore, M.A., Scheufele, D.A., Corley, E.A. (2013). Another (methodological) look at knowledge gaps and the Internet's potential for closing them. *Public Understanding of Science*, 0, 1-9.

- Campbell, A., Gurin, G., & Miller, W.E. (1954). *The Voter Decides*. Evanston, Illinois: Row, Peterson and Company.
- Clarke, H.D., Kornberg, A., & Scotto, T.J. (2008). *Making Political Choices: Canada and the United States*. Toronto: University of Toronto Press.
- Clemence, M., Gilby, N., Shah, J., Swiecicka, J., Warren, D., Smith, P., Johnson, J., Hoolahan, F., D'Souza, J. (2013). *Wellcome Trust Monitor Wave 2: Tracking public views on science, biomedical research and science education*. London: Wellcome Trust.
- Craig, C., Niemi, R.G., Silver, G.E. (1990). Political Efficacy and Trust: A Report on the NES Pilot Study Items. *Political Behavior*, 12, 289-314.
- Eveland JR., W.P., Marton, K., Seo, M. (2004). Moving Beyond "Just the Facts": The Influence of Online News on the Content and Structure of Public Affairs Knowledge. *Communication Research*, 31, 82-108.
- Eveland JR., W.P.a.S., D.A. (2000). Connecting News Media Use with Gaps in Knowledge and Participation. *Political Communication*, 17, 215-237.
- Haustein, S., Peters, I., Sugimoto, C.R., Thelwall, M., & Lariviere, V. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, 65, 656-669.
- Ho, S.S. (2012). The Knowledge Gap Hypothesis in Singapore: The Roles of Socioeconomic Status, Mass Media, and Interpersonal Discussion on Public Knowledge of the H1N1 Flu Pandemic. *Mass Communication and Society*, 15, 695-717.
- Kahan, D.M., Braman, D., Slovic, P., Gastil, J., Cohen, G. (2009). Cultural Cognition of the Risks and Benefits of Nanotechnology. *Nature Nanotechnology*, 4, 87-91.

- Kahan, D.M., Wittlin, M., Braman, D., Slovic, P., Peters, E., Ouellette, L.L., Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature, Climate Change*, 732-735.
- Kaid, L.L., McKinney, M.S., Tedesco, J.C. (2007). Political Information Efficacy and Young Voters. *American Behavioral Scientists*, 50, 1093-1111.
- Kenski, K.A., Stroud, N.J. (2006). Connections Between Internet Use and Political Efficacy, Knowledge, and Participation. *Journal of Broadcasting & Electronic Media*, 50, 173-192.
- Kim, S.-H. (2008). Testing the Knowledge Gap Hypothesis in South Korea: Traditional News Media, the Internet, and Political Learning. *International Journal of Public Opinion Research*, 20, 193-210.
- Kruikemeier, S., van Noort, G., Vliegenthart, R., de Vreese, C.H. (2014). Unraveling the effects of active and passive forms of political Internet use: Does it affect citizens' political involvement? *New Media & Society*, 16, 903-920.
- Ladwig, P., Dalrymple, K.E., Brossard, D., Scheufele, D.A., & Corley, E.A. (2012). Perceived familiarity or factual knowledge? Comparing operationalizations of scientific understanding. *Science and Public Policy*, 39, 761-774.
- Lee, C.-j. (2009). The Role of Internet Engagement in the Health-knowledge Gap. *J Broadcast Electron Media*, 53, 365-382.
- Leifeld, P. (2013). texreg: Conversion of Statistical Model Output in R to LaTeX and HTML Tables. *Journal of Statistical Software*, 55, 1-24.
- Liu, H., Priest, S. (2009). Understanding public support for stem cell research: media communication, interpersonal communication and trust in key actors. *Public Understanding of Science*, 18, 704-718.

- Masic, I., & Milinovic, K. (2012). On-line Biomedical Databases-The Best Source for Quick Search of the Scientific Information in the Biomedicine. *Acta Informatica Medica*, 20, 72-84.
- Mejlgaard, N. (2009). The trajectory of scientific citizenship in Denmark: Changing balances between public competence and public participation. *Science and Public Policy*, 36, 483 – 496.
- Mejlgaard, N., & Stares, S. (2010). Participation and competence as joint components in a cross-national analysis of scientific citizenship. *Public Understanding of Science*, 19, 545 – 561.
- Mejlgaard, N., & Stares, S. (2013). Performed and preferred participation in science and technology across Europe: Exploring and alternative idea of "democratic deficit". *Public Understanding of Science*, 22, 660-673.
- Mellenbergh, G.J. (1996). Measurement Precision in Test Score and Item Response Models. *Psychological Methods*, 1, 293-299.
- Miller, J.D. (2001). Who Is Using the Web for Science and Health Information? *Science Communication*, 22, 256-273.
- Mirisola, A., & Seta, L. (2016). pequod: Moderated Rgression package. <https://CRAN.R-project.org/package=pequod>.
- Moeller, J., de Vreese, C., Esser, F., Kunz, R. (2014). Pathway to Political Participation: The Influence of Online and Offline News Media on Internal Efficacy and Turnout of First-Time Voters. *American Behavioral Scientists*, 58, 689-700.
- Morrell, M.E. (2003). Survey and Experimental Evidence for a Reliable and Valid Measure of Internal Political Efficacy. *Public Opinion Quarterly*, 67, 589-602.



Muthén, L. K., Muthén, B. O. (1998-2010). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Niemi, R.G., Craig, S.C., & Mattei, F. (1991). Measuring Internal Political Efficacy in the 1988 National Election Study. *American Political Science Review*, 85, 1407-1413.

Nisbet, M.C., & Scheufele, D.A. (2004). Political talk as a catalyst for online citizenship. *Journalism & Mass Communication Quarterly*, 81, 877-896.

Scheufele, D.A., & Nisbet, M.C. (2002). Being a citizen online - New opportunities and dead ends. *Harvard International Journal of Press-Politics*, 7, 55-75.

Shapin, S. (2008). *The Scientific Life*. Chicago: The University of Chicago Press.

ONS (2010). *Standard Occupational Classification 2010*, 2010. Office for National Statistics.

ONS (2012). *Statistical Bulletin: Internet Access - Households and Individuals, 2012*. Office for National Statistics.

ONS (2015). *Statistical Bulletin: Internet Access - Households and Individuals:2015*. Office for National Statistics.

ONS (2016). *Statistical Bulletin: Internet users in the UK: 2016*. Office for National Statistics.

Su, Y.L., Akin, H., Brossard, D., Scheufele, D.A., Xenos, M.A. (2015). Science News Consumption Patterns and Their Implications for Public Understanding of Science. *Journalism & Mass Communication Quarterly*, 92, 597-616.

Su, Y.L., Cacciatore, M., Scheufele, D.A., Brossard, D., Xenos, M.A. (2014). Inequalities in Scientific Understanding: Differentiating Between Factual and Perceived Knowledge Gaps. *Science Communication*, 36, 352-378.

Tedesco, J.C. (2007). Examining Internet Interactivity Effects on Young Adult Political Information Efficacy. *American Behavioral Scientists*, 50, 1183-1194.

Tedesco, J.C. (2011). Political Information Efficacy and Internet Effects in the 2008 U.S. Presidential Election. *American Behavioral Scientists*, 55, 696-713.

Tewksbury, D., Hals, M.L., Bibart, A. (2008). The Efficacy of News Browsing: The Relationship of News Consumption Style to Social and Political Efficacy. *Journalism and Mass Communication Quarterly*, 85, 257-272.

Tichenor, P.J., Donohue, G.A., & Olien, C.N. (1970). Mass Media Flow and Differential Growth in Knowledge. *Public Opinion Quarterly*, 34, 159-170.

## **4. “Excuse me, doctor...” What drives people’s confidence to participate in medical decision-making?**

### **4.1. Introduction**

#### 4.1.1. The doctor-patient relationship

The rapport between doctors and patients has been the subject of a vast array of medical, philosophical, literary, and, amongst others, sociological treatises since Hippocrates. The manner in which doctors and their patients interact has been undergoing a deep transformation in the last fifty years, however, brought about by the proliferation of the notion of patient autonomy. This concept of patient autonomy has, in fact, occupied a central locus in sociology of medicine debates since the 1970s (Britten, 2001). In medical ethics, the moral notion of autonomy was subsumed as one of the *prima facie* principles meant to guide the doctor-patient interaction (Goold and Lipkin, 1999). In the 1980s, the World Health Organisation declared that patient participation in healthcare is beyond desirable, namely, it is a technical, social, and economic requisite (Charles, Gafni, Whelan, 1999).

Moreover, Chin (2002) illustrates the transition of medical ethics principles inscribed in the American Medical Association (AMA) ethical code between 1847 and 1990. In 1847, the AMA titled their ethical code “Obligations of patients to their physicians”. In 1990, this was drastically reformed to “Fundamental Elements of the Patient-Physician Relationship”. The first describes a paternalistic model of patient-doctor interaction, while the second marks an important step towards a shared decision-making model of patient-doctor interaction (Charles, Gafni, Whelan, 1999; Gafni, Charles, and Whelan, 1998; Deber et al., 2007; DeWalt, 2007).

In the UK, the Royal Pharmaceutical Society of Great Britain published a document in 1997, called ‘From compliance to concordance’, which reconceptualised the patient-doctor relationship as one between equals (cited in Britten, 2001). Numerous socio-economic and psychosocial factors are postulated to influence individuals’ desire for involvement in the decisions regarding their healthcare. Studies have predominantly reported that younger age, female sex, higher education, being white, and less serious illness are the main drivers of a greater impetus to participate in medical decision-making. Be that as it may, these socio-demographic and health influences only make up part of the story. Psychological traits, health literacy, self-efficacy, and trust in medics have all been identified as determinants of the public partaking in their healthcare (Fortune et al., 2016; Sarkar, Fisher, Schillinger, 2006; Flynn and Smith, 2007; Ishikawa and Yano, 2008; for a meta-analysis see Benbassat et al., 1998). It is of great interest to disentangle the seemingly opposite loci of impact of efficacy and trust, as well as any socio-economic differences in how these value predispositions operate. To this end, the present study uses conditional inference tree models to examine the roles of efficacy and trust in influencing willingness to participate in medical decision-making. Modelling them together and allowing a computer algorithm to split them according to what combinations are meaningful in predicting participation might provide further insight into, for instance, whether high trust and high efficacy together are similar or different to low trust and low efficacy. The role of long-term illness or disability, knowing someone in the medical profession, education, gender, ethnicity, and age is also explored. In the interest of a more democratic form of medicine and clinical practice, as well as a more comprehensive public policy, it is important to further understand the factors that affect citizens’ willingness to be more proactive in situations where their wellbeing is at stake.

#### 4.1.2. Models of patient-physician interaction

But before delving into a more in-depth review of findings in relation to what makes people confident to participate in a clinical situation, an overview of the principle tenets of the main models of physician-patient interaction is deemed important. It serves to carve a theoretical and empirical path to the examination of individual confidence to challenge their main healthcare provider and, often the first port of call, the General Practitioner, as well as the socio-economic factors affecting it. The latter constitutes the scope of the present study. Prior to the 1990s, the paternalistic model was the dominant doctor-patient model of communication and decision-making, with the physician assuming the dominant role. The physician's control over the clinical situation stemmed from knowledge, power, status, and sometimes gender asymmetry. The recurring motive, however, was that of doctor beneficence (Pellegrino, 1994).

Charles, Gafni, and Whelan (1999) identify four main principles underlying the deference to medico-professional authority. The first is physicians' medical proficiency and the fact that they are up-to-date with clinical thinking and findings. The second belief is that they apply this information consistently and equally across clinical situations and patients to select the one best treatment available. Thus, physicians are not just the most well-informed; they consistently use this information to select the best course of clinical action for each of their patients. The third principle states that clinicians' expertise in medical matters enables them to distinguish between different treatment options and select the single best course of clinical action for each of their patients. The fourth and last postulation states that physicians are invested in each decision they make, and that the patient's best interest is accorded primordial importance. The latter was, in fact, legitimised through the implementation of codes of ethics, such as the AMA code, which bound doctors to put their patients' best interests first (Charles, Gafni, Whelan, 1997).

It becomes apparent that in the paternalistic model, patients' own wishes, concerns, and their whole persona, are left out of the decision-making process. Furthermore, individuals' socio-economic circumstances, attitudes, self-efficacy, how much they trust doctors, as well as their fears, which have been shown to affect people's health (Guadagnoli and Ward, 1998), are simply ignored. New and transformative ideals regarding the inclusion of some or all of these factors in the patient-doctor relationship took the form of new models of communication and decision-making, proposed by scholars, medical professionals, and regulators. Two new models briefly discussed here are the shared decision-making model and the informed model. More specifically, both of these models are designed to remedy the first problem discussed, namely involving patients in the medical conversation. They give the patient a voice in a clinical situation by opening up a two-way dialogue. The decision is then either shared or left to the patient (Deber et al., 2007).

The shared decision-making model was proposed in order to diminish the knowledge and power asymmetry between patients and physicians. The physician possesses greater technical knowledge in most cases, and whilst that remains unchallenged, its communication to the patient is what becomes central to the new model. In other words, the physician should share their information with the patient, present treatment options, and then they reach a consensus on the best course of action together (Charles, Gafni, Whelan, 1999; Deber et al., 2007; DeWalt, 2007). The informed model takes the degree of patient participation to a new level, being postulated as almost the reverse of the paternalistic model of medicine that dominated for centuries. After imparting their medical knowledge on the specifics of the clinical situation to the patient, as well as the treatment options available, the physician takes a step back and allows the patient to arrive at their own conclusions and choose the best treatment for themselves (Deber et al., 2007, DeWalt, 2007).

#### 4.1.3. Preferences for participation and socio-economic differences

The above alterations to the conceptualisation of the patient-physician dyad happened in the legal, political, sociological, and medico-regulatory realms. What patients themselves, or, indeed, public opinion, make of the disintegration of the paternalistic model in the Anglo-American context might be an entirely different matter. Studies report that individuals do not want to have a great deal of involvement in the decisions regarding their treatments (Ahmad et al., 2006; Deber et al., 2007; DeWalt, 2007). The latter has made some scholars more reluctant to support models that were concocted as alternatives to the paternalistic model (Guadagnoli and Ward, 1998; Sanders et al., 2013).

Deber et al. (2007) administered the Problem-Solving Decision-Making Scale (PSDMS) to subgroups of outpatients from a large teaching hospital in Ontario, Canada. PSDMS is an instrument designed to measure patients' preferred role in both discussing the health issue and making a decision regarding the course of treatment. Patients are presented with different clinical scenarios, such as a general health status check, chest pain, etc. They found that very few desired full autonomy (i.e. only 1.2% out of over 1,000 when presented with a chest pain clinical scenario). Most respondents preferred to make the decision jointly with their doctor (65.1%), and a passive role came next (34.1%).

Kraetchmer et al. (2004) report similar findings from an urban teaching hospital in Canada. Only 1.2% out of the 606 respondents preferred a fully autonomous role in a chest pain scenario, 48.7% people exhibiting a preference towards shared decision-making, with a small majority of 50.1% preferring a passive role. However, it is not just in potentially life-threatening situations that individuals do not want full autonomy over the clinical situation or decision regarding their treatment. When presented with the current health condition scenario, respondents' preferences followed the same pattern as above, with very small

upward deviations in their preferences for autonomy. For instance, 2.9% desired full autonomy instead of 1.2%. Interestingly, in the current health status scenario, 67.3% of patients desired shared decision making, and only 29.6% wanted a passive role. It seems that in more serious situations, these patients were more willing to defer to a medical authority, whereas in a non-threatening general check-up scenario, they claimed some of the authority back. Vertinsky (1974) and Haug and Lavin (1981) (both cited in Guadagnoli and Ward, 1998) report similar findings on small samples of healthy respondents.

In the UK, Robinson and Thomson (2001) conduct focus groups and report that individuals may not understand the importance of being more involved in their healthcare decisions. Moreover, they make an argument for amplified information transmission from physicians to their patients, which would lead to more confidence to partake. The realisation of the importance of their participation would be actualised post factum. Numerous other studies in medical sociology, medicine, and public health literature have reported findings following a similar pattern (see for example Arora and McHorney, 2000; Benbassat et al., 1998; Charles, Gafni, Whelan, 1999; Deber et al., 2007; DeWalt, 2007; Fortune Shotwell, Buccelato, and Moran, 2016,).

This study sets out to overcome some of the limitations of the findings discussed. These limitations include some methodological issues, such as small and unrepresentative samples often sourced from quite specific populations (for example patients in a certain clinic), mostly urban settings in the US and Canada. There were only a handful of studies identified that are based on a representative sample. Levinson, Kao, Kuby, and Thisted (2005) used the General Social Survey and found that nearly all respondents preferred to be asked their opinion and to be offered choices. About half wanted to leave the final decision to their



physicians, and 44% preferred to receive information from doctors rather than looking for it themselves.

#### 4.1.4. Efficacy, Trust, and Participation in medical decision-making

Socio-economic factors do not paint the full picture of what motivates, or indeed deters, people from actively partaking in decisions regarding their health (Benbassat et al., 1998). Beyond socio-economic influences, there are important psychosocial mechanisms underpinning attitudes towards taking part in healthcare decision processes. Amongst these, self-efficacy and trust in doctors as information sources are of central interest in the present study. Seeing as participatory ideas have an element of underlying confidence, self-efficacy in reference to biomedical science and research seems particularly relevant when attempting to explain predilections towards medical decision-making. Self-efficacy has the potential to determine the amount of effort individuals will put into an activity/behaviour as well as their persistence in the face of hurdles and stressful situations (Bandura, 1977; Pellegrino, 1994).

Multiple past studies have found that people who feel more efficacious have more favourable attitudes towards partaking in medical decision-making as well as augmented participation rates (Stretcher, McEnvoy, Becker, and Rosenstock, 1986; Janz et al., 2004; Belcher et al., 2005). The mechanism postulated is in line with Bandura's conceptualisation of efficacy, namely that increased feelings of confidence in one's competence lead to greater effort put into certain activities and behaviours. Moreover, the political science concept of internal efficacy is what the construct of biomedical self-efficacy has been based on in this PhD thesis (see for example Craig, Niemi, Silver, 1990; Morrell, 2003; Niemi, Craig, Mattei, 1991; Kaid, McKinney, Tedesco, 2007). Going further than self-reported familiarity with biomedical terms (referred to as self-reported knowledge in the Public Understanding of Science literature), in this chapter efficacy also englobes a participatory aspect, namely

confidence to discuss science heard or seen on the news. This is similar to some of the participatory questions used by Mejlgaard and Stares (2010) to measure scientific citizenship.

Initially, public participation in medical decision-making was also modelled as a component of efficacy, in an attempt to simulate the efficacy measure used in Chapter 3. However, a Confirmatory Factor Analysis model revealed that participation in healthcare decisions does not cohere well with the measures of self-reported knowledge, or those of the usefulness and application of science, even though they are correlated. This may be because it does not ask respondents to relate to medicine as an entity, as the participation component used in Chapter 3 did. Seeing as the self-reported familiarity questions also refer to biomedicine as an entity (as opposed to a personal experience of it – i.e. having stem cells collected), relating them to a question about a potential personal clinical experience seems to not work. Willingness to take part in public consultations regarding medical research may have worked well with the self-assessment questions in Chapter 3 because individuals are asked to rapport their confidence to bigger issues, which is more similar to internal political efficacy. A clinical scenario may trigger different thought and emotional processes, where respondents can use their sense of efficacy about biomedicine and science in general to inform their behaviour. Thus, conceptually and empirically, participation in healthcare decisions represents a different construct. Ergo, it makes sense to study the impact of biomedical efficacy on this type of participation.

All in all, one of the aims of this paper is to study the relationship between confidence in one's own abilities and potential behaviour during clinical practice. Confidence about knowing the meaning of a biomedical term, confidence to discuss science heard on the news, and reporting to find it useful in everyday situations, might or might not translate into

confidence in a clinical practice situation where the person's health might be at stake, hypothetical though it may be.

Moreover, a related concept, namely health literacy, has also been discovered to be positively related to increased participation rates (Ishikawa and Yano, 2008). Using the Wisconsin Longitudinal Graduate survey, a population-based cohort of older adults aged 63 – 66, Flynn and Smith (2007) investigate the relationship between personality traits, which are related to self-efficacy (Bandura, 1977), and a preference for participation in medical decisions. Openness to experience, decreased neuroticism and agreeableness were related to preferences for a more active decision-making style. These psychological traits are not measured here. However, any findings regarding high levels of confidence to participate would also tell us a little about their underlying psychological characteristics.

Returning to biomedical self-efficacy, the mechanisms relating efficacy to participation identified in the literature review are also expected to be at work here. The first hypothesis thus becomes:

**H1:** People with higher levels of biomedical self-efficacy will show more willingness to participate in medical decision-making.

The other potential psychosocial influence explored here is trust in medics and other healthcare professionals as information sources. The issue of who the public place their trust in for information regarding their healthcare is crucial in a society that sees fast-advancing technologies and medical therapies (Anderson et al., 2012, 2008). The dynamic of the relationship between trust and participatory ideas is conceptually and empirically distinct to that connecting efficacy to participation. Unlike confidence in one's own abilities, which prompts a higher desire for participation, trust in doctors has been found to work more in a

similar vein to deference to scientific authority. Given the breadth of medical knowledge, most citizens don't have the time to keep up with the latest developments. Therefore, when making decisions about their healthcare they rely on cognitive shortcuts such as trust. Higher levels of trust appear to be forged by the older and more highly educated citizens deferring more to scientists/specialists (Brossard and Nisbet, 2006).

The principles underlying deference to scientific authority are in line with the cognitive miser theory, whereby citizens face a choice of who to trust rather than formulating well-informed opinions regarding the benefits and risks of certain scientific technologies, or courses of treatment in this case (Anderson et al., 2012). The main tenet of deference to scientific authority is that the scientific establishment has already determined what is right or wrong and non-scientists should not form their own views regarding scientific controversies (Brossard and Nisbet, 2006; Anderson et al., 2012;). Anderson et al. (2012) postulate that the main dissimilarity of deference from trust is its stability and longevity, trust being subject to change. Stability and longevity resonate with the paternalistic model of patient-physician interaction, which was prevalent for decades. Because of its longevity, paternalistic clinical practice is bound to have institutionalised a culture of deference to doctors in terms of medical decisions. This is probably especially true amongst the socio-demographic groups identified as wanting to be less involved, given their lower socio-economic characteristics and the power dynamic in a clinical scenario. The tide has only recently started to change and public levels of willingness to partake in decisions concerning their health are still low.

Finally, trust in doctors as information sources here is posited to be quite similar to deference to scientific authority. Findings in the health literature also suggest that, in general, trust in doctors and trust in doctors as information sources operates in a similar way to deference. In other words, individuals who tend to exhibit higher trust levels want to be less involved in

the decision-making regarding their healthcare, and have higher adherence to treatment plans. It could thus be argued that they defer to their physicians (Pearson and Raeke, 2000; Belcher, 2005; Hesse et al., 2005; Trachtenberg, Duga, and Hall, 2005; DeWalt, 2007; Fortune et al., 2016).

This leads to the second hypothesis:

**H2:** People with higher levels of trust in doctors as information sources will show less willingness to participate in medical decision-making.

In terms of socio-demographic differences, studies have shown younger age, female sex, higher education, being white, and less serious illness to be related to more positive attitudes towards participation in healthcare decisions (Trachtenberg, Dugan, Hall, 2005; Fortune et al., 2016). Chronic conditions seem to influence preferences differently to the acute medical scenarios discussed in the literature review. The author is not aware of a similar population-based study in the UK examining socio-economic differences in participation preferences. One of the benefits of the current study is that it is based on a representative sample of the UK population, which can potentially offer insight into preferences of the public at large, rather than just specific patient groups. It can thus help shed light on broader population patterns in preferences for participation in medical decision-making, and potentially provide some evidence for policy makers. Given the scarcity of population-based studies, and the somewhat mixed results of the impact of socio-economic factors, this analysis can be regarded as an exercise in both prediction and classification of preferences for participation. Thus, I pose the following research question:

**RQ1:** What are the interrelationships amongst efficacy, trust, and socio-economic factors that lead to different individual participation preferences?

## 4.2. Method

### 4.2.1. Data

In order to test these hypotheses, conditional inference tree models were fitted using the adult sample from Wave III of the Wellcome Trust Monitor Survey. The Wellcome Trust Monitor is a triennial survey of adults and young people based on a representative probability sample of the UK population. The main areas covered are interest in, knowledge of, and attitudes to medical research, with a focus on biomedicine. Wellcome Monitor III took place in 2015, and interviews were conducted with 1,524 adults (aged 18+). The response rate for Monitor III was 51.4% for the adult sample (Huskinson et al., 2016).

### 4.2.2. Measures

The outcome variable, *willingness/confidence to participate in medical decision-making/challenge GPs*, was measured with a single item. Respondents were given the following scenario and then were asked to say how much the person in the scenario was, or was not, like them: ‘*When he sees a medical professional, he likes to ask questions to make sure the medical professional hasn’t made a mistake or missed something, and feels confident querying or challenging the medical professional’s conclusions. How much is this person like you?*’. The response options were as follows: 1. Very much like me; 2. Like me; 3. Somewhat like me; 4. A little like me; 5. Not at all like me; 6. SPONTANEOUS: Can’t say/ it depends. The last option was dropped as it bears no substantive meaning for these analyses. The variable was recoded as a new ordered variable with the following categories: 1. Very much like me; 2. Like/Somewhat like me; 3. A little/not at all like me. It was treated as an ordered factor in R. As a note: willingness to participate/participation/confidence to challenge GPs or to participate are used interchangeably throughout this chapter.

The *biomedical self-efficacy* construct was measured by combining responses to five questions. Two of the questions ask respondents to assess their understanding of two key terms in biomedical research, namely ‘Genetically Modified’ and ‘Antibiotic Resistance’. The options presented to them are: 1. Very good; 2. Good; 3. Some understanding; 4. Have heard the term but have little understanding of what it means; and 5. Have not heard the term. These were reverse coded so that: ‘Have not heard to term’ became ‘0’; ‘Have heard the term but have little understanding of what it means’ was coded as ‘1’; ‘Some understanding’ was recoded to ‘2’; ‘Good’ became ‘3’; and ‘Very good’ was recoded to ‘4’. The resulting summated scale ranged from 2 to 21. It was recoded to range from 0 to 19.

The next two questions tap into a more general self-reported understanding of science news stories and confidence to discuss these. The first asks: ‘Thinking of the stories about science you see or hear in the news, which of the statements on this card would you say best describes you?’. Respondents are then presented with four options: 1. I usually understand what they are talking about; 2. I sometimes understand what they are talking about; 3. I usually do not understand what they are talking about; 4. (SPONTANEOUS: I don’t see or hear science new stories). The fourth option was dropped from these analyses as it doesn’t say anything about their confidence regarding their level of understanding; it seems to simply suggest that they may not have heard anything related to science in the news, or may not be willing to cognitively engage enough to remember. This decision was further motivated by the fact that there were only 21 individuals who answered in this way. Response options 1 through 3 were reverse coded from negative to positive so that ‘3’ became ‘0’, ‘2’ became ‘1’, and ‘1’ became ‘2’. This question is then followed by: ‘And in general, how confident would you say you feel discussing these news stories about science with other people?’. Participants were then asked to choose from the following options: 1. Very confident; 2. Confident; 3. Not very confident. 4. Not at all confident. Answers were

again reverse coded so that '4' became '0', '3' became '1', '2' remained '2', and '1' became '3'. The last question taps into respondents' subjective assessments of the usefulness of science in their everyday lives. They then have to choose from the following options: 1. Very useful. 2. Fairly useful. 3. Not very useful. 4. Not at all useful. 5. I don't know enough about science for it to be useful. Again, the scale was reverse coded to follow a negative to positive pattern. Option '5' was recoded to '0', '4' to '1', '3' to '2', '2' to '1', and '1' to '4'.

Initially, an exploratory correlation matrix was run in order to explore the degree of relatedness amongst these variables, which ranged from 0.25 to 0.58. Other variables asking about confidence to use their knowledge to inform aspects of their lives, such as making a decision to make an appointment to see a doctor, were initially included, but the correlations with the other variables were too weak. A further Confirmatory Factor Analysis treating the final selected variables as ordinal measures was conducted, and the hypothesized one factor structure fit the data well (CFI= 0.98; TLI= 0.97; RMSEA=0.04). Results from two CFA analyses are presented in the Appendix. The first considers the possibility that willingness to participate in healthcare decisions is part of efficacy. However, this item does not load on the efficacy factor. The second CFA analysis yields a one-factor solution for the remaining five questions described above. For parsimony, a summated scale was built by simply adding answers to all five questions, which demonstrated strong internal consistency (Cronbach's alpha= 0.81). This is used in all subsequent analyses.

*Trust in doctors as information sources* were measured with a single 5-point item. Respondents were asked to rate their trust in five science actors to provide 'accurate and reliable information about medical research.' The first of the series of five actors were 'doctors, nurses, and other medical practitioners'. Respondents then had to say whether they had: 1. Complete trust; 2. A great deal of trust; 3. Some trust; 4. Very little trust; 5. No trust



at all. The scale was reversed to range from '1', for 'No trust at all', to '5', for 'Complete trust'. The variable was treated as an ordered factor in R.

In line with findings outlined in the previous section, gender, age, education, ethnicity, knowing a medical professional, and having a disability or long-term illness were introduced as control measures. Gender was used as a binary variable. Ethnicity was used as a dummy variable representing white and non-white individuals (whites are the reference category). Age was recoded from a continuous variable ranging from 18 to 96 to three binary variables representing three age categories, namely a young group aged 18 to 34, a middle-aged group aged 35 to 54, and an older group aged over 55, respectively. In addition, a dummy variable denoting whether respondents know or do not know a medical professional was also included in the model. The rationale for this is that individuals might be more comfortable asserting themselves in front of doctors if they at least know someone in that profession.

Education was measured with three dummy variables coded from a variable which measures respondents' highest educational qualification. The initial categories were: 1. Postgraduate degree; 2. First degree; 3. Higher education below degree; 4. A levels/ GCSEs/O levels; 5. Level 1 qualifications; 6. No qualifications. The first three categories were collapsed to form a new dummy variable representing those who had gone into higher education. Category 4 became a dummy variable. Categories 5 and 6 were collapsed and recoded into a binary variable. 'Higher education' was used as the reference category.

In order to measure the presence of a disability or long-term illness, respondents were asked the following question: *'Do you yourself or someone close to you, have a serious long term illness or medical condition? Please include any conditions which have lasted or are likely to last at least 12 months and which, without treatment, would limit one's ability to carry out normal day-to day activities.'* The response options were: 1. Yes, respondent; 2. Yes, family

**Table 4.1 Descriptive statistics**

<b>Variable name</b>	<b>Mean/%</b>	<b>S.D.</b>
<b>Participation/confidence to challenge medical professional:</b>		
<i>Very much like me (v)</i>	30	
<i>Like/Somewhat like me (l/s)</i>	42	
<i>A little/Not at all like me (li/n)</i>	28	
<b>Biomedical self-efficacy</b>	12.1	2.94
<b>Trust:</b>		
1.No trust at all	2	
2.Very little trust	5	
3.Some trust	30	
4.A great deal of trust	43	
5.Complete trust	20	
<b>Age dummies:</b>		
18-34 years old	24	
35-54 years old	33	
Over 55 years old	43	
<b>Education level:</b>		
Degree+	40	
A level/ GCSE/O level	35	
Level 1/ no qualifications	25	
<b>Gender (male)</b>	44	
<b>Has disability/long-term illness</b>	24	
<b>White</b>	91	

member; 3. Yes, close friend; 4. No, no one. A dummy variable was created. It was coded '1' for those who mentioned having a disability or long-term illness themselves, and '0' if they mentioned any of the other three options, namely close family, friends, or no one. I have not been able to find any studies that examine a potential link between a family member, or friend, being ill and desire for involvement in decisions about one's own health. One of the objectives of this study is to use a large scale survey to shed some light on the association

between individual socio-economic characteristics and participation in medical decision-making, which has thus far only been investigated using small patient groups from hospitals or clinics. Therefore, individual disability or illness, which has been examined in those studies, is of central interest.

Distribution frequencies of all variables used in the regression analyses are shown in Table 4.1. Means are given for continuous measures and percentages represent the proportion of respondents who fall into each category. Plots showing bivariate relationships between each explanatory variable and the outcome are shown in the Appendix to this Chapter.

#### 4.2.3. Analytical Strategy

All data management and data analysis were carried out in R 3.3.2/ RStudio 1.0.44. Conditional inference trees modelling participation as an ordinal variable were fitted to the data. First, ordered logistic regression was attempted, but the proportional odds assumptions did not hold. A conditional inference tree is a data mining technique, nonparametric and algorithmic in nature, which uses binary recursive partitioning to model regressions of a set of covariates on the outcome of interest. It can estimate an ordinal regression tree without the proportional odds assumption. Unlike the mixed logit models shown in the Appendix, it also retains the ordinal nature of the outcome variable. The algorithm tests if any covariates are associated with the given outcome, and picks the variable that has the strongest association with it. It then makes a binary split in this predictor, dividing the dataset into two or more subsets, depending on the measurement level. The algorithm is recursive because the first two steps are repeated for each subset until there are no variables that are associated with the outcome at the pre-defined level of statistical significance (I have chosen 0.05 as the cut-off). The results can be visualized as a tree structure with binary splits forming ‘branches’ and ‘leaves’. Respondents are assigned to ‘nodes’, or classes, based on the subsets they fall into

during the final split, at the end of the recursive partitioning process. The algorithm uses permutation to perform the partitioning, which is similar to bootstrapping. It draws random samples from the subsets in order to perform the significance level tests (Hothorn et al, 2006a; Hothorn, Hornik and Zeileis, 2006b; Hothorn and Zeileis, 2015).

Packages ‘party’ and ‘partykit’ are the only packages that can implement ordinal conditional inference trees in R (Hothorn, Hornik and Zeileis, 2006b; Hothorn and Zeileis, 2015). Conditional inference trees are also more robust than traditional decision/regression trees as they deal with the problem of overfitting and use significance level tests instead of information measures. Random forests of individual trees can also be fitted, and they work by extracting a user-specified number of trees through resampling strategies. Variable importance measures can then be used to assess each covariate’s contribution. For ordinal outcomes they measure by how much the probability of a given response would decrease if a certain covariate was to be left out of the analysis (Hothorn et al., 2006a). Random forests of multiple conditional inference trees were also estimated. Two thousand random trees were estimated. Random forests can improve the prediction accuracy compared to estimating a single tree, but are currently somewhat experimental in nature, and thus are not fully understood. Therefore, a random forest is only estimated to check if it improves the prediction accuracy and to check the importance of each variable in the model (Hothorn et al., 2006a). Model fit statistics were obtained using package ‘caret’ (Kuhn, 2008).

### **4.3. Results**

The tree plot in figure 4.1 below illustrates the relationships between the covariates and the classification of individuals into nodes, based on their characteristics and probabilities of responding a certain way. The circles represent internal nodes, and at the bottom of the tree there are terminal nodes. The latter show a final classification of individuals based on the

paths amongst the internal nodes and their final probabilities of response to the outcome variable. Nodes are also referred to as classes. 'Node' and 'class' shall be used interchangeably from this point forth. The height of the bars in the terminal classes corresponds to the probability of response to the outcome variable, participation. The categories were relabelled so that the names could be easily read on the tree plot. 'Like/somewhat like me' became 'l/s'; 'A little/ not at all like me' became 'li/n'; 'Very much like me' became 'v'. The tree had a moderate prediction accuracy of 49% and a fair inter-rater agreement ( $\kappa=0.21$ ; this is similar to a correlation coefficient and it ranges from '0' to '1') (Landis and Koch, 1975), These fit statistics tell us how well the predictors work together, and how helpful these categories, that people are assigned to based on the most meaningful combinations of their characteristics that can be related to the outcome, are in predicting attitudes to participation. The prediction accuracy for the random forest with 2000 random trees was not better than this. Prediction rates and kappa coefficients were obtained by comparing predicted to true/observed values following the recommended procedure in the 'party' package vignette (Hothorn et al., 2006a).

From figure 4.1 it emerges that there are eight terminal nodes containing respondents with various characteristics, that are represented in the inner nodes. These nodes are arrived at by following the interrelationships illustrated by the branches (paths) that go from internal node to internal node. The classes vary in sizes somewhat, although in general the differences are not great. Individuals in class 5 have the highest probability of not wanting to be involved in decisions made by their GPs. This class is very small, however, comprising of only 13 people. The largest class found, namely class 6 ( $n=336$ ) is next in terms of it being composed of respondents who have little or no confidence to take part in the medical decision-making process. Individuals in class 13 ( $n=169$ ) show more favourability towards being involved, having the highest proclivity to say it is like or somewhat like them to

challenge and query a GP's decisions. The most active and confident of respondents fall into class 15 (n=270).

In a way, the splits can also be regarded as similar to interacting variables. The tree plot reveals the different combinations of efficacy, trust, and socio-economic factors that are significant in understanding preferences for, and confidence to participate in, medical decision-making. Understanding these interrelationships helps answer **RQ1**. Respondents in node 1 who score 11 or lower on efficacy, and who know medical professionals, fall into terminal node 3. This latter class is characterised by higher probabilities of saying that querying a GP is 'like or somewhat like' them, followed by a 'little or not at all like' them, with 'very much like' them coming last. The path from node 2 to node 4 is characterised by not knowing medical professionals. Respondents who do not know medical professionals, and score 3 or lower, out of 19, on the efficacy scale, are finally allocated to node 5 which only has respondents who have low confidence to participate in decisions regarding their health. If, however, respondents in node 4 score over 3 on efficacy, they are allocated to node 6. The latter class is also defined by higher probabilities of low confidence to participate, and has a gradient in confidence, with slightly fewer respondents having medium confidence to participate, and even fewer exhibiting high confidence.

We can observe more interactions by following the paths amongst nodes 1, 7, 8, 10, 12 to terminal nodes 9, 11, 13, 14, and 15. Following the path from node 1 to node 7 are respondents scoring above 11 on the efficacy scale. Node 7 then splits respondents between those that fall above 15 on the efficacy scale, and those that fall at or below 15, respectively. The first split then follows a path to class 8, representing trust. As a reminder, trust categories represent the following levels: 1. No trust at all; 2. Very little trust; 3. Some trust; 4. A great deal of trust; 5. Complete trust. They appear as numbers on the graph due to

readability issues (the words are too large to appear on the graph). It is treated as an ordinal variable in R.

If individuals have either no trust or some trust in doctors and other medical practitioners, they are split further and allotted to terminal node 9. This class is characterised by a probability of being very confident in querying GP diagnoses that is only very slightly higher than moderate confidence (saying 'like/ somewhat like me'). Individuals allocated to this final class have significantly lower probabilities of saying it is only a 'little or not at all like' them to challenge GPs decisions regarding their healthcare.

Thus, an absence of trust in doctors or only having some trust leads to respondents not deferring to them for medical decisions, which is in consonance with previous findings and provides evidence to support the second hypothesis. Oddly, some trust appears to work the same as no trust, and very little trust is grouped with high levels of trust. A potential explanation is that those who respond that they have some trust are more cognitively disengaged than those who say they have no trust at all. The other potential explanation is that those with very little trust at node 8 they are further split at node 10 into those who have a long-term illness and end up in terminal class 11, which is similar to class 9 in relation to participatory attitudes. This, however, cannot be checked with this model.

Going back to node 8 representing the split for trust, individuals who say they have very little trust, a great deal of trust, or complete trust are further split at node 10 into those who have a long-term illness or disability and those who do not. The first are then allocated to final class 11 who are most likely to say it is like or somewhat like them to query their General Practitioner. This is followed by a slightly lower probability of saying it is 'very much like' them, and a much lower probability of having little or no confidence to challenge their GPs.

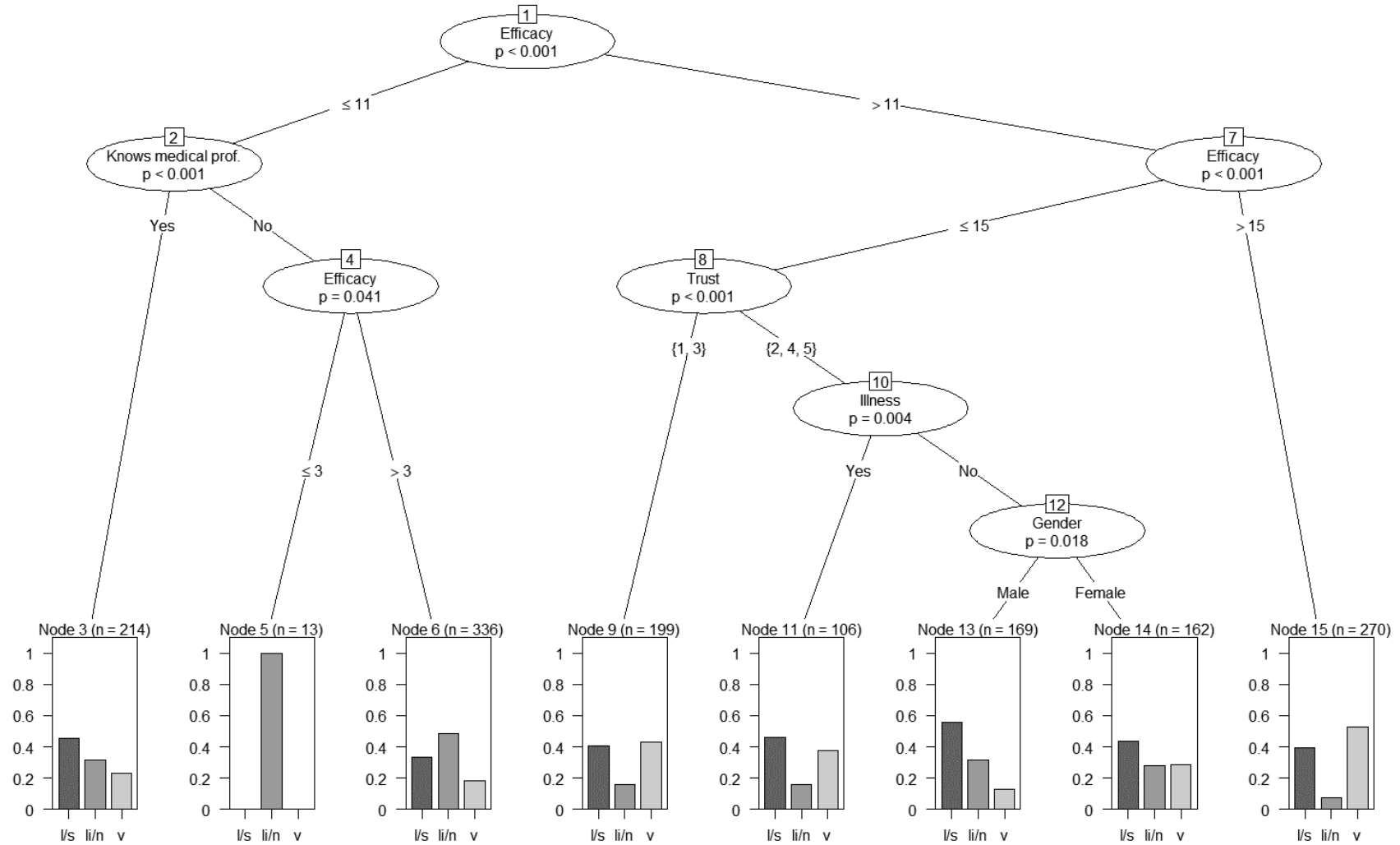
The second split for node 10 represents those who do not have a long-term illness or disability, and who are allocated to node 12, where a further split is made based on gender. Thus, those who do not have a long-term illness and who are male are allocated to final class 13. The latter is characterised by high probabilities of saying it is only 'like or somewhat like' them to query GPs. There also seems to be a stepwise decrease in probability, with individuals in this class having the lowest probability of being highly confident, compared to the other two levels of confidence. However, overall, this still have higher probabilities of medium willingness to participate.

Returning to node 12 representing gender, the path for women sees them apportioned to terminal node 14, also defined by a higher probability of being confident or somewhat confident to participate in medical decision-making. This is lower than for men in class 13. The probabilities of either having no confidence or very much confidence appear equal. A relatively high level of efficacy (12-15 out of 19), little or great levels of trust, no long-term disability, and being male leads to a greater chance of having low or medium confidence to challenge medics, rather than high confidence. Compared to men, women with the same characteristics have higher probabilities of being very confident, in line with other findings.

To understand class 15 the path back to node 7 needs to be followed. The second split from node 7 is based on respondents scoring over 15 on the efficacy scale, who are then assigned to node 15. It seems that individuals who are very self-efficacious with regards to their knowledge of biomedicine also have the highest probability of being very confident in challenging medical diagnoses. This is in line with previous findings and lends support to the first hypothesis.



**Figure 4.1 Conditional inference tree predicting participation**

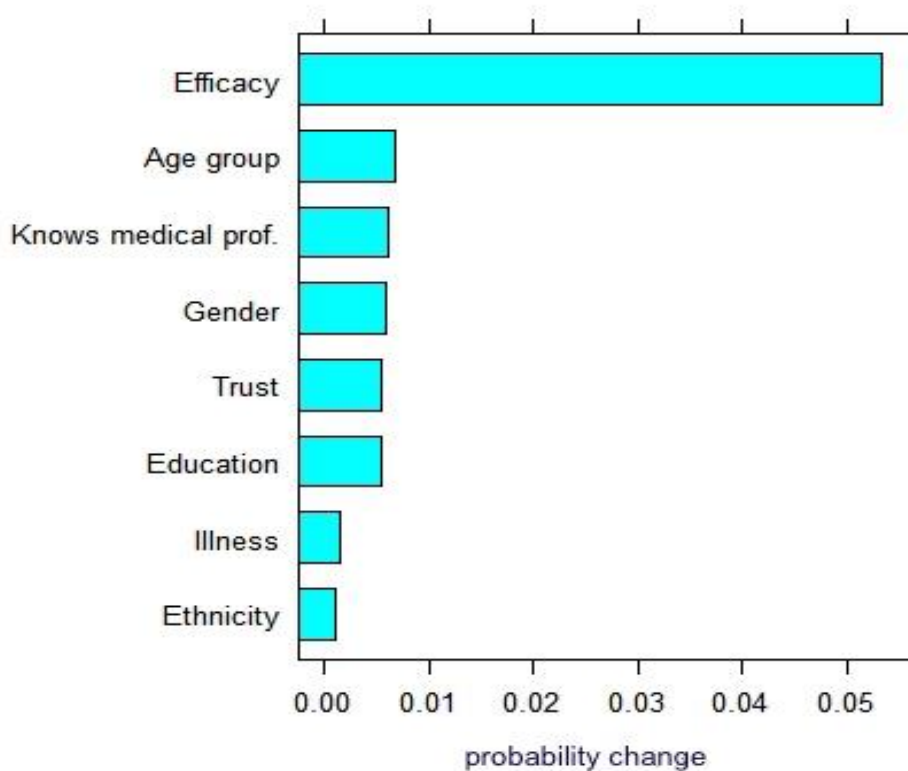


Overall, class 15 is the most self-efficacious and has the highest propensity to be willing to partake in the medical decision-making process. Thus, with this sort of confidence and self-efficacy, an argument can be made that this class might possess a favourable predisposition towards an informed decision-making model.

Figure 4.2 shows the importance of each variable included in the random forest model. Tree plots from random forests cannot be plotted, but they are used to assess the change in the probability of the outcome if an individual predictor were removed from the analysis (seen on the x axis).

Efficacy seems to be the main driver of participatory attitudes. Age, knowing a medical professional, gender, trust, and education seem to have a similar influence in predicting participatory attitudes.

**Figure 4.2 Variable importance plot**



Age and education were not statistically significant, however, therefore they do not appear in the tree plot. Having a long-standing illness of disability and ethnicity have a similar impact too, but the latter is not statistically significant.

#### **4.4. Conclusion**

So far, most studies investigating public willingness and confidence to participate in medical decision-making have been conducted on small unrepresentative samples. Moreover, most of these studies have been carried out in the US and Canada. Very few have attempted to delve into the complexities of British public attitudes towards their involvement in the process of making medical decisions, or how this fits in with the greater debates regarding patient-physician interactions. The present study addresses the latter and it uses a representative sample of the UK population, thus overcoming a lot of the shortcomings of most past studies. Understanding which, if any, patient-physician interaction model the British public might show preference for is a very important step in potentially informing the remodelling of the healthcare system in the UK into a more democratic establishment.

Two hypotheses and a research question were posed. The first hypothesis posited that higher self-efficacy is linked to higher willingness and self-assurance to question a physician's decision. The second hypothesis advanced that higher deference to medical authority, in the form of a high level of trust in doctors, nurses, and other medical practitioners as information sources, leads to lower willingness to participate. The research question investigates the interrelationships of socio-economic factors, efficacy, trust, and participation willingness. All findings discussed below answer this research question.

A conditional inference tree was fitted in R using package ‘party’, which estimates an ordinal regression tree without imposing the restrictive proportional odds assumption, which is a feature of the traditional technique (Hothorn, Hornik and Zeileis, 2006b).

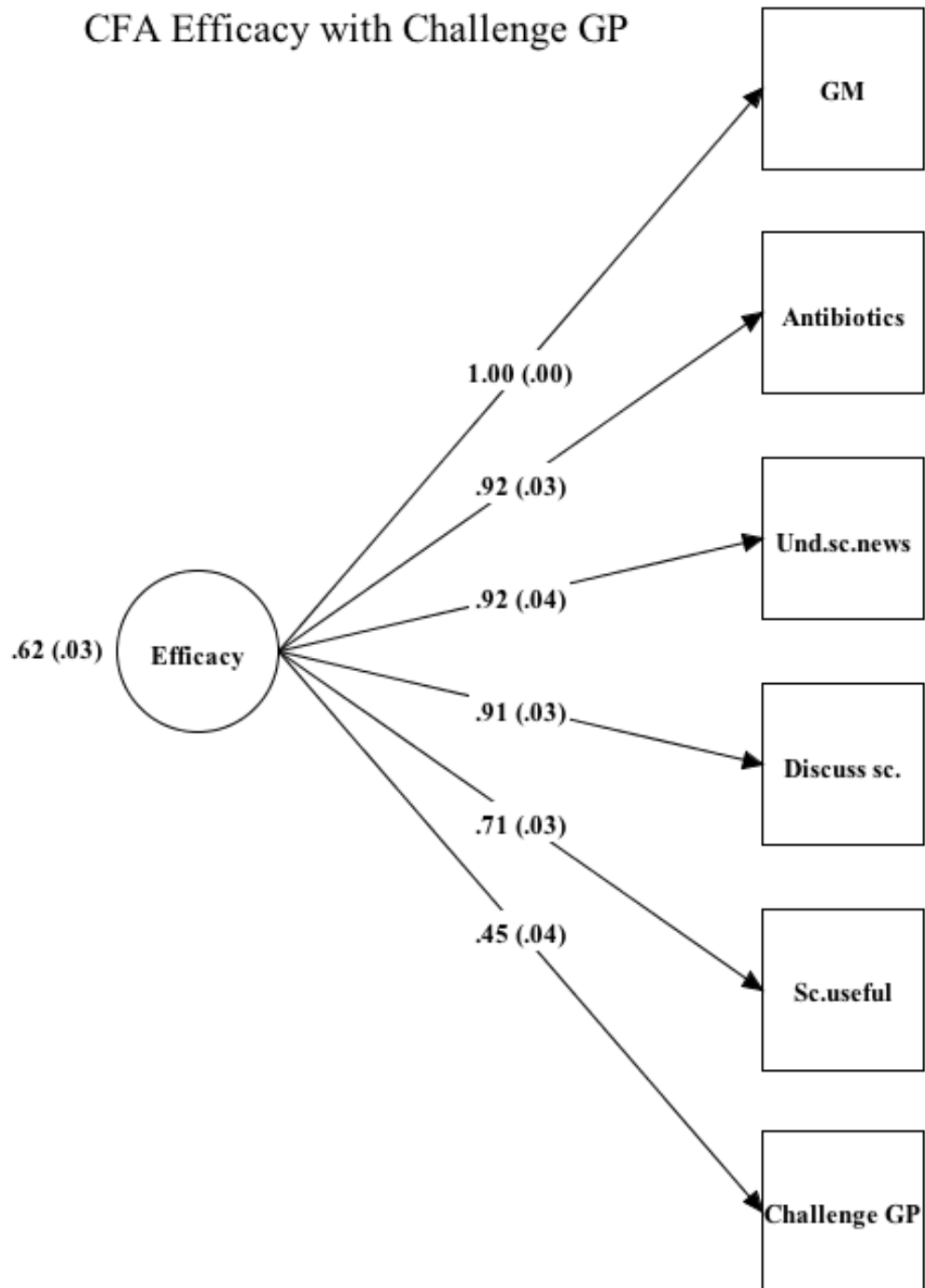
A total of eight final groups categorising individuals based on the combinations of the predictors were found. Overall, most of the groups have moderate desire and confidence to participate in medical decisions, with most classes having higher probabilities of saying it is ‘like or somewhat like’ them to question a GP’s decisions. The individuals in the class who have the highest probability of saying it is ‘somewhat like’ them or ‘like’ them to be active in a clinical scenario are male, have either little trust or a lot of trust, have no long-term illness or disability, and are very confident in their abilities to understand and discuss biomedicine and science. Thus, there is evidence to support the first hypothesis seeing that higher self-efficacy is associated with higher propensities towards more active involvement on the respondents’ part. The second hypothesis also receives partial support in that some of these individuals who are highly willing and confident to participate have little trust in doctors. This reinforces findings from past studies (Pearson and Raeke, 2000; Belcher, 2005; Hesse et al., 2005; Trachtenberg et al., 2005; DeWalt, 2007; Fortune et al., 2016).

The group who seem to be the least disposed to participate (who have the highest probability of not wanting to be involved out of all the groups) have very low self-efficacy and do not know medical professionals personally. However, there are only thirteen individuals in this group, so it is hard to make any further inferences about them. The largest class found is also the class that is characterised by the next highest propensity to have little or no confidence to take part in decisions regarding their health. They are also not self-efficacious when it comes to their knowledge of biomedicine. Thus, it could be argued that these respondents would adhere more to the paternalistic model of patient-doctor contact. Two groups of individuals

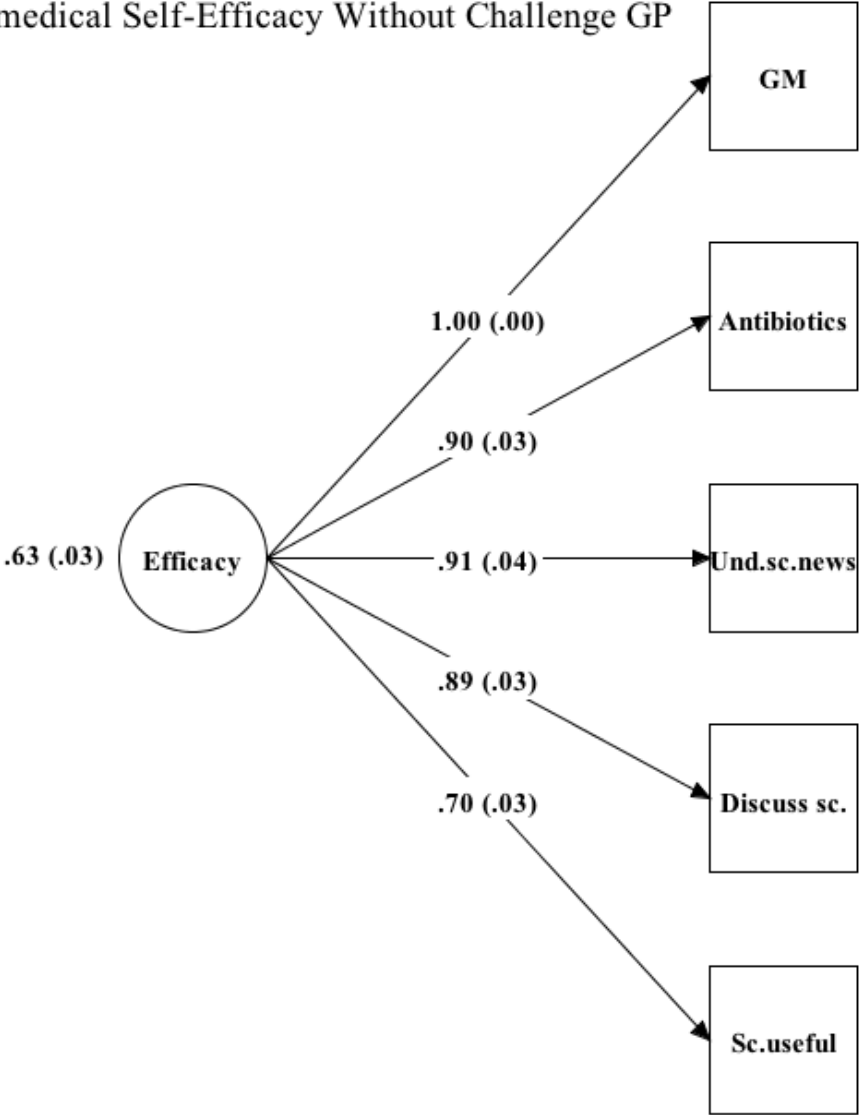
showed higher chances of very high confidence to be involved. The first only exhibits a marginally higher likelihood of a high degree of confidence, is very self-efficacious, and does not defer to medical authority. The second has extremely high levels of self-efficacy and an even higher probability of participation. This, again, is in line with other findings from past studies, and it may indicate tendencies towards a preference for an informed model (Stretcher, McEnvoy, Becker, Rosenstock, 1986; Janz et al., 2004; Belcher et al., 2005).

Further thoughts on the theoretical implications of my findings can be found in Chapter 5.

## 4.5 Appendix

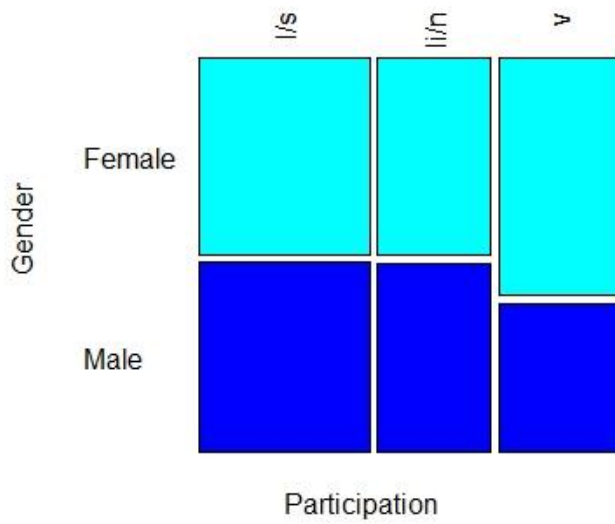


CFA Biomedical Self-Efficacy Without Challenge GP

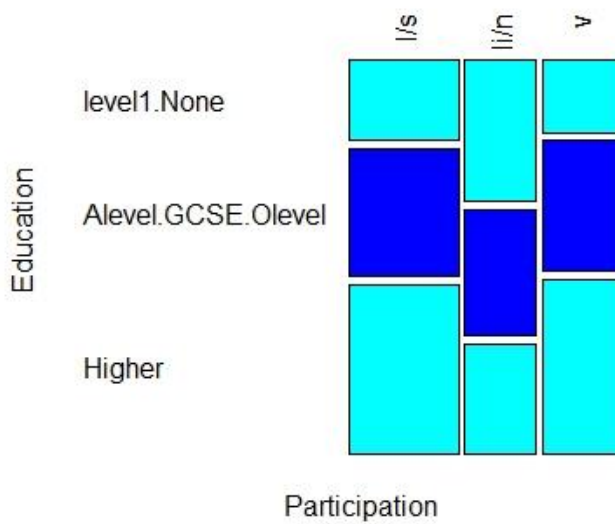


**Bivariate relationships between each predictor and the outcome:**

**Participation by gender**

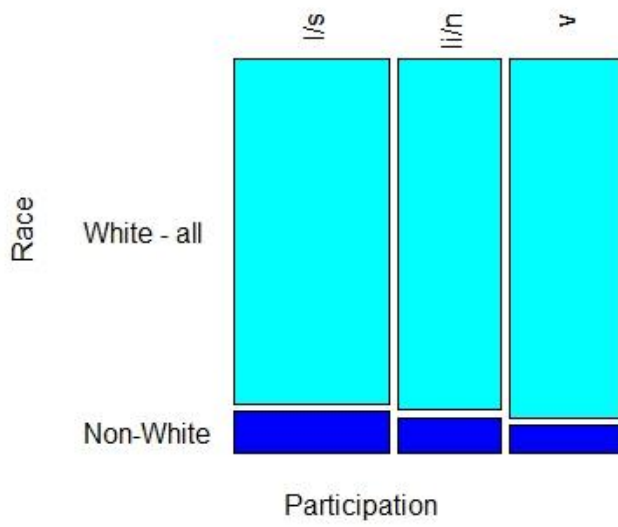


**Participation by education**

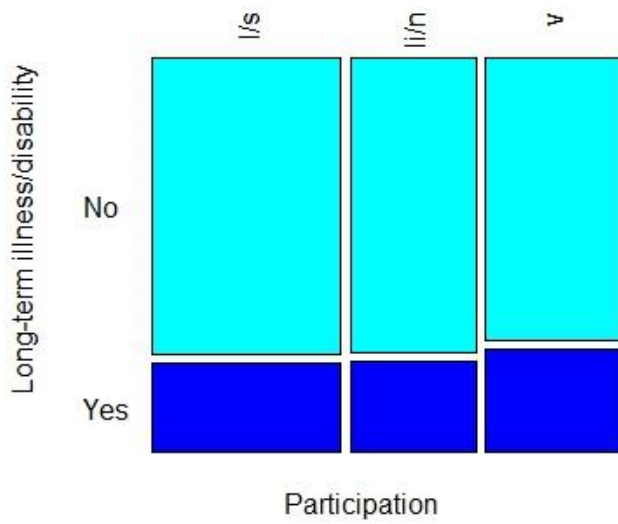




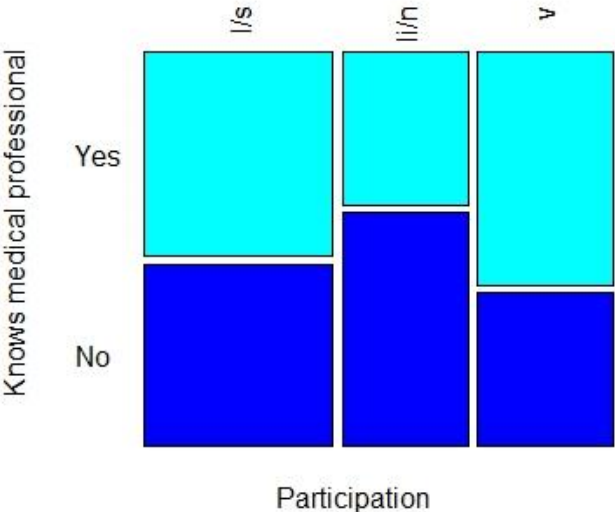
### Participation by race



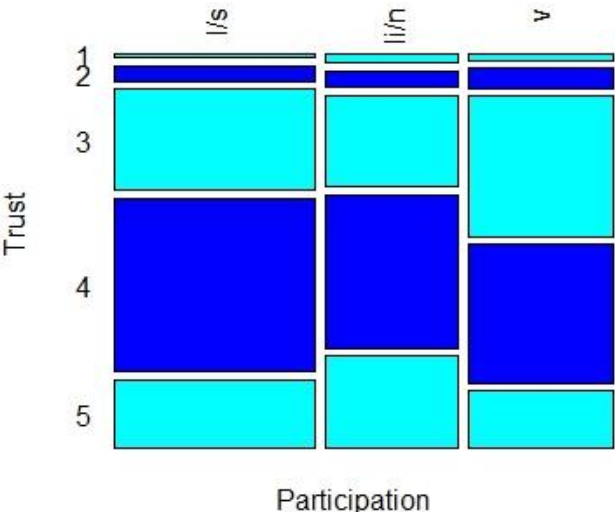
### Participation by long-term illness



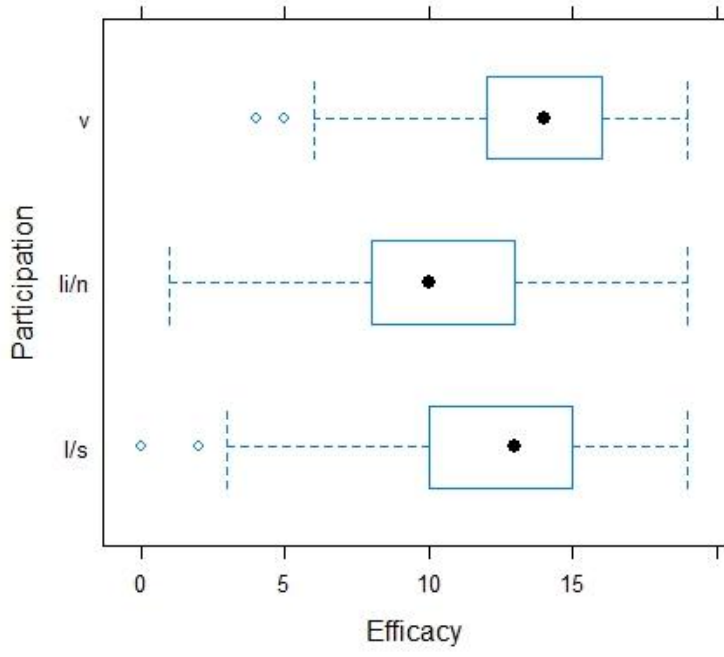
**Participation by knows medical professional**



**Participation by trust in doctors**



### Participation by efficacy



### **A Mixed Logit Analysis:**

A somewhat different version of this analysis was also expressed as mixed logit models for the same reasons conditional inference trees were estimated. First, ordered logistic regression was attempted, but the proportional odds assumptions did not hold. Multinomial logistic regression was considered next, but the independence of irrelevant alternatives (IIA) assumption seemed neither plausible in this case, nor did it succeed to pass the Hausman MacFadden test. The mixed logit model relaxes the IIA assumption and allows for random parameters to be introduced, thus allowing estimates to vary from individual to individual. This is in contrast to multinomial logit models, where they are fixed. Coefficients are considered random variables, and random numbers are drawn from their joint conditional distribution. The unconditional probability is then calculated as the average of the conditional probabilities. Because the probabilities are integrals that have closed form, the estimation of mixed logit models is done by simulation. Halton sequence numbers have been found to be the most efficient, and are employed in these models (McFadden and Train, 2000; Croissant, 2013). Models were run with 500 and 1000 random draws, on two different seeds, namely 20 and 40. Results were almost identical, so the coefficients presented below are from the models run with 500 random draws, seed 20. ‘Like/very much like me’ was chosen as the reference category because it is the largest category. Random intercepts for ‘a little/not at all like me’ and ‘somewhat like me’ were modelled initially to test whether there was evidence of individual heterogeneity in preferences for participation. The distribution of these random parameters was first specified as normal and then lognormal. However, in both cases, the standard deviations for the two parameters were not statistically significant from 0, therefore, they were then removed from the model. All other parameters remained the same after the removal of the random intercepts. The simulations were carried on as explained above. Interactions with educational variables were attempted in these models. Variables

were introduced in what appeared to be the logical causal order. The outcome variable was recoded as follows: ‘like me’ and ‘very much like me’ were collapsed into a category; the middle option ‘somewhat like me’ was retained; ‘a little like me’ and ‘not at all like me’ were collapsed. No claims to causality are made seeing that these models use cross-sectional data to only test for associational relationships. Tables were created using package ‘stargazer’ (Marek, 2015).

Trust and efficacy were modelled separately as these models have slightly different hypotheses. Additional hypotheses:

**H3:** Having low education, will negatively moderate the impact of efficacy on participation willingness.

**H4:** Higher education will positively moderate the impact of trust on participation willingness.

Models 1 show results for efficacy and trust, respectively. Models 2 add socio-demographics. Models 3 include the interaction terms between science training and the efficacy and trust variables, respectively. Models 4 present results for the two-way interactions between highest educational qualification and efficacy and trust, respectively. Figure 4.3 illustrates the significant interactions between efficacy and science training. Figure 4.4 portrays the interaction between trust and education.

**Table 4.2 Mixed logit- efficacy models**

	Dependent variable:			
	Participation(ref: Like/Very much like me)			
	(1)	(2)	(3)	(4)
A little/Not at all:Intercept	9.063***	5.107***	3.210***	6.100***
Somewhat:Intercept	1.693***	1.637***	0.789	2.843***
A little/Not at all:Efficacy	0.790***	0.799***	0.834***	0.789***
Somewhat:Efficacy	0.899***	0.891***	0.952***	0.854***
A little/Not at all:Male		1.555***	1.559***	1.560***
Somewhat:Male		1.261***	1.268***	1.277***
A little/Not at all:35-54		1.025***	1.021***	1.026***
Somewhat:35-54		1.021***	1.016***	1.029***
A little/Not at all:18-34		1.360***	1.353***	1.353***
Somewhat:18-34		1.142***	1.133***	1.118***
A little/Not at all:Level1/None		1.643***	1.629***	1.009
Somewhat:Level1/None		1.070***	1.071***	0.250
A little/Not at all:Alevel/GCSE/Olevel		1.080***	1.025***	0.970
Somewhat:Alevel/GCSE/Olevel		0.924***	0.860***	0.702
A little/Not at all:Illness/disability		0.770***	0.777***	0.768***
Somewhat:Illness/disability		0.810***	0.818***	0.807***
A little/Not at all:Science Training		1.131***	3.218***	1.142***
Somewhat:Science Training		1.129***	5.129***	1.164***
A little/Not at all:Efficacy X Science Training			0.918***	
Somewhat:Efficacy X Science Training			0.884***	
A little/Not at all:Efficacy X Level1/None				1.041
Somewhat:Efficacy X Level1/None				1.136
A little/Not at all:Efficacy X Alevel/GCSE/Olevel				1.007
Somewhat:Efficacy X Alevel/GCSE/Olevel				1.017
LR Test	173.165*** (df= 6)	198.877*** (df = 20)	208.145*** (df = 22)	205.097*** (df = 24)

**N=1465; Note:\*p<0.1; \*\*p<0.05; \*\*\*p<0.01**

**Table 4.3 Mixed logit- trust models**

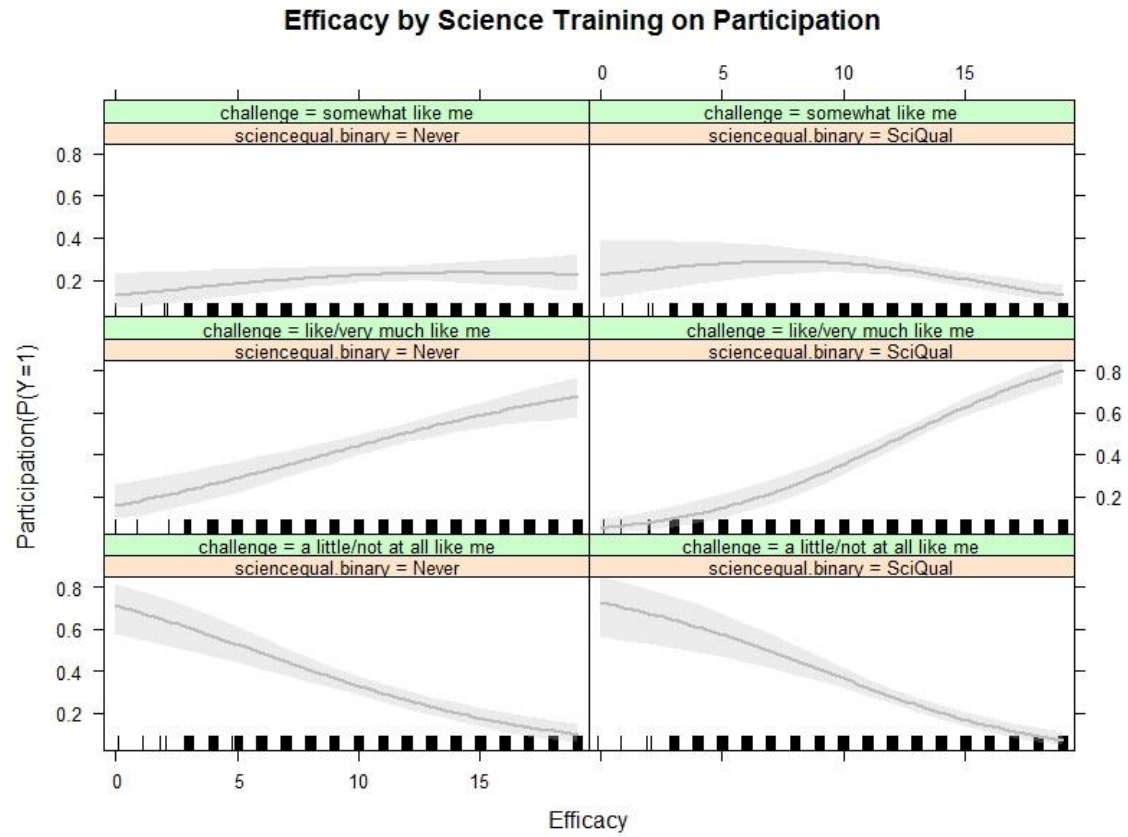
	Dependent variable:			
	Participation (ref:Like/Very much like)			
	(1)	(2)	(3)	(4)
A little/Not at all:Intercept	0.600***	0.281	0.252	0.386
Somewhat:Intercept	0.358	0.274	0.234	0.389
A little/Not at all:Trust.L	1.076***	1.335***	1.656***	0.546
Somewhat:Trust.L	2.083***	2.229***	2.646***	0.799
A little/Not at all:Trust.Q	1.783***	1.385***	0.979***	5.435***
Somewhat:Trust.Q	0.946**	0.864**	0.445	2.547***
A little/Not at all:Trust.C	0.941***	0.862***	0.850***	0.465
Somewhat:Trust.C	0.934***	0.912***	0.962**	0.533
A little/Not at all:Trust^4	0.955***	0.949***	0.903***	0.982**
Somewhat:Trust4	1.061***	1.068***	0.952***	1.216***
A little/Not at all:Male		1.288***	1.305***	1.301***
Somewhat:Male		1.139***	1.160***	1.162***
A little/Not at all:35-54		1.089***	1.087***	1.093***
Somewhat:35-54		1.063***	1.062***	1.058***
A little/Not at all:18-34		1.817***	1.820***	1.801***
Somewhat:18-34		1.325***	1.318***	1.321***
A little/Not at all:Level1/None		3.348***	3.400***	2.391***
Somewhat:level1/None		1.581***	1.641***	0.999**
A little/Not at all:Alevel/GCSE/Olevel		1.568***	1.554***	0.948**
Somewhat:Alevel.GCSE.Olevel		1.131***	1.121***	0.795*
A little/Not at all:Illness/Disability		0.829***	0.822***	0.843***
Somewhat:Illness/Disability		0.843***	0.833***	0.845***
A little/Not at all:Science Training		0.844***	1.087***	0.848***

Somewhat:Science Training		0.969***	1.312***	0.973***
A little/Not at all:Trust.L X Science Training			0.584	
Somewhat:Trust.L X Science Training			0.660	
A little/Not at all:Trust.Q X Science Training			2.117	
Somewhat:Trust.Q X Science Training			3.621	
A little/Not at all:Trust.C X Science Training			0.960	
Somewhat:Trust.C X Science Training			0.824	
A little/Not at all:Trust^4 X Science Training			1.109	
Somewhat:Trust^4 X Science Training			1.252	
A little/Not at all:Trust.L X Level1/None				2.172**
Somewhat:Trust.L X Level1/None				3.255***
A little/Not at all:Trust.Q X Level1/None				0.186
Somewhat:Trust.Q X Level1/None				0.201
A little/Not at all:Trust.C X Level1/None				1.659**
Somewhat:Trust.C X Level1/None				1.474*
A little/Not at all:Trust^4 X Level1/None				1.053**
Somewhat:Trust^4 X Level1/None				0.698
A little/Not at all:Trust.L X Alevel/GCSE/Olevel				4.640***
Somewhat:Trust.L X Alevel/GCSE/Olevel				3.127**
A little/Not at all:Trust.Q X Alevel/GCSE/Olevel				0.137
Somewhat:Trust.Q X Alevel/GCSE/Olevel				0.345
A little/Not at all:Trust.C X Alevel/GCSE/Olevel				2.917***
Somewhat:Trust.C X Alevel/GCSE/Olevel				2.158**
A little/Not at all:Trust^4 X Alevel/GCSE/Olevel				0.920*
Somewhat:Trust^4 X Alevel/GCSE/Olevel				1.093**
LR Test	20.430*	93.602***	102.438***	110.358***
	(df = 12)	(df = 26)	(df = 34)	(df = 42)

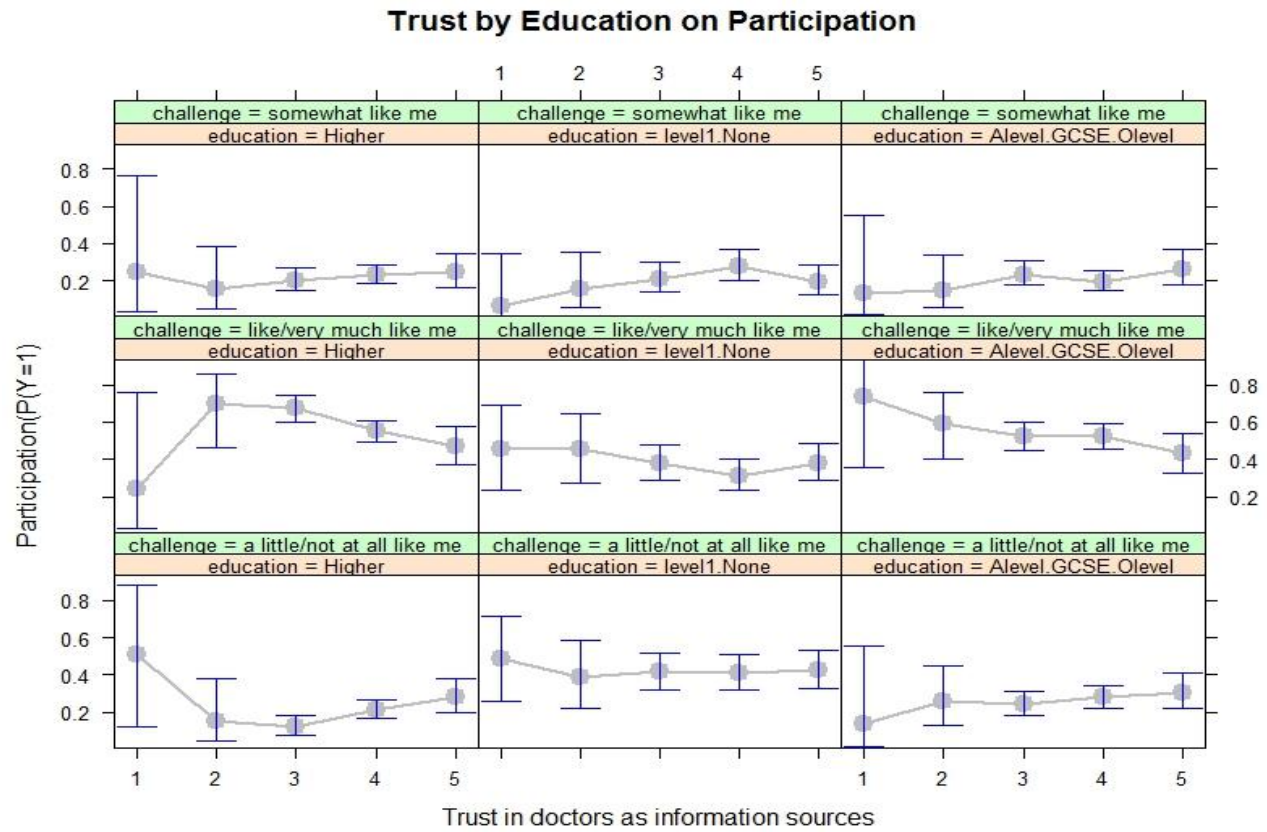
**N=1465; Note:\*p<0.1; \*\*p<0.05; \*\*\*p<0.01**



Figure 4.3



**Figure 4.4**



## References

- Ahmad, F., Hudak, P., Bercovitz, K., Hollenberg, E., & Levinson, W. (2006). Are Physicians Read for Patients With Internet-Based Health Information? *Journal of Medical Internet Research*, 8, e22.
- Anderson, A.A., Scheufele, D.A., Brossard, D., & Corley, E.A. (2012). The Role of Media and Deference to Scientific Authority in Cultivating Trust in Sources of Information about Emerging Technologies. *International Journal of Public Opinion Research*, 24, 225-237.
- Arora, N.K., & McHorney, C.A. (2000). Patient Preferences for Medical Decision Making: Who Really Wants to Participate? *Medical Care*, 38, 335-341.
- Bandura, A. (1977). Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, 84, 191-215.
- Belcher, V., N., Fried, T.R., Agostini, J.V., & Tinetti, M.E. (2006). Views of Older Adults on Patient Participation in Medication-related Decision Making. *Journal of General Internal Medicine*, 21, 298–303.
- Benbassat, J., Pilpel, D., & Tidhar, M. (1998). Patients' preferences for participation in clinical decision making: a review of published surveys. *Behavioral Medicine*, 24, 81-88.
- Britten, N. (2001). Prescribing and the defence of clinical autonomy. *Sociology of Health & Illness*, 23, 478-496.
- Brossard, D., & Nisbet, M.C. (2006). Deference to scientific authority among a low information public: Understanding U.S. opinion. *International Journal of Public Opinion Research*, 19, 24-52.

Charles, C., Gafni, A., & Whelan, T. (1997). Shared Decision-Making in the Medical Encounter: What Does it Mean? (Or it Takes at Least Two to Tango). *Social Science & Medicine*, 44, 681-692.

Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Social Science & Medicine*, 49, 651-661.

Chin, J. (2002). Doctor-patient Relationship: from Medical Paternalism to Enhances Autonomy. *Singapore Medical Journal*, 43, 152-155.

Craig, C., Niemi, R.G., Silver, G.E. (1990). Political Efficacy and Trust: A Report on the NES Pilot Study Items. *Political Behavior*, 12, 289-314.

Croissant, Y. (2013). mlogit: multinomial logit model. <https://CRAN.R-project.org/package=mlogit>

Deber, R.B., Kraetschmer, N., Urowitz, S., & Sharpe, N. (2007). Do people want to be autonomous patients? Preferred roles in treatment decision-making in several patient populations. *Health Expectations*, 10, 248-258.

DeWalt, D.A., Boone, R.S., & Pignone, M.P. (2007). Literacy and Its Relationship With Self-efficacy, Trust, and Participation in Medical Decision Making. *American Journal of Health Behavior*, S27-S35.

Flynn, K.E., & Smith, M.A. (2007). Personality and Health Care Decision-Making Style. *Journals of Gerontology. Series B: Psychological Sciences and Social Sciences*, 62, 261–267.

Fortune, E.E., Shotwell, J.J., Buccellato, K., & Moran, E. (2016). Factors predicting desired autonomy in medical decisions: Risk-taking and gambling behaviors. *Health Psychology Open*, January-June 2016, 1-10.

Gafni, A., Charles, C., & Whelan, T. (1998). The Physician-Patient Encounter: The Physician as a Perfect Agent for the Patient Versus the Informed Treatment Decision-Making Model. *Social Science & Medicine*, 47, 347-354.

Goold, S.D., & Lipkin, M. (1999). The doctor-patient relationship. Challenges, opportunities, and strategies. *Journal of General Internal Medicine*, 14, S26-S33.

Guadagnoli, E., & Ward, P. (1998). Patient participation in decision-making. *Social Science & Medicine*, 47, 239-339.

Hesse, B.W., Nelson, D.E., Kreps, G.L., Croyle, R.T., Neeraj, A.K., Rimer, B.K., et al. (2005). Trust and Sources of Health Information: The Impact of the Internet and Its Implications for Health Care Providers: Findings From the First Health Information National Trends Survey. *Archives of Internal Medicine*, 165, 2618-2624.

Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. (2006a). Survival Ensembles. *Biostatistics*, 7, 355-373.

Hothorn, T., Hornik, K., & Zeileis, A. (2006b). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15, 651-674.

Hothorn, T., & Zeileis, A. (2015). partykit: A Modular Toolkit for Recursive Partytioning in R. *Journal of Machine Learning Research*, 16, 3905-3909.

Huskinson, T., Gilby, N., Evans, H., Stevens, J., Tipping, S. (2016). Wellcome Trust Monitor Wave 3: Tracking public views on science, biomedical research and science education. London: Wellcome Trust.

Ishikawa, H., & Yano, E. (2008). Patient Health literacy and participation in the health-care process. *Health Expectations*, 11, 113-122.

Janz, N.K., Wren, P.A., Copeland, L., A., Lowery, J.C., Goldfarb, S.L., & Wilkins, E.G. (2004). Patient-Physician Concordance: Preferences, Perceptions, and Factors Influencing the Breast Cancer Surgical Decision *Journal of Clinical Oncology*, 22, 3091-3098.

Kaid, L.L., McKinney, M.S., Tedesco, J.C. (2007). Political Information Efficacy and Young Voters. *American Behavioral Scientists*, 50, 1093-1111.

Kraetschmer, N., Sharpe, N., Urowitz, S., & Deber, R.B. (2004). How does trust affect patient preferences for participation in decision-making? *Health Expectations*, 7, 317–326.

Kuhn, M. (2008). Caret package. *Journal of Statistical Software*, 20, 1-26.

Landis, J.R., & Koch, G.G. (1975). A review of statistical methods in the analysis of data arising from observer reliability studies (Part I). *Statistica Neerlandica*, 29, 101-123.

Levinson, W., Kao, A., Kuby, A., & Thisted, R.A. (2005). Not All Patients Want to Participate in Decision Making: A National Study of Public Preferences *Journal of General Internal Medicine*, 20, 531–535.

Marek, H. (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2. <http://CRAN.R-project.org/package=stargazer>.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete choice. *Journal of Applied Econometrics*, 15, 447-470.

Morrell, M.E. (2003). Survey and Experimental Evidence for a Reliable and Valid Measure of Internal Political Efficacy. *Public Opinion Quarterly*, 67, 589-602.

NICE (2013). Patient and Public Involvement Policy, pp. 1-15. <https://www.nice.org.uk/Media/Default/About/what-we-do/Into-practice/Field-team/field-team-report-2013-14-v7.pdf>: National Institute for Health and Care Excellence. [Accessed January 20, 2017]

Niemi, R.G., Craig, S.C., & Mattei, F. (1991). Measuring Internal Political Efficacy in the 1988 National Election Study. *American Political Science Review*, 85, 1407-1413.

Pearson, S.D., & Raeke, L.H. (2000). Patients' Trust in Physicians: Many Theories, Few Measures, and Little Data. *Journal of General Internal Medicine*, 15, 509-513.

Pellegrino, E., D. (1994). Patient and Physician Autonomy: Conflicting Rights and Obligations in the Physician-Patient Relationship. *Journal of Contemporary Health Law & Policy*, 10, 47-68.

Robinson, A., & Thomson, R. (2001). Variability in patient preferences for participating in medical decision making: implication for the use of decision support tools. *BMJ Quality & Safety*, 10, 34-38.

Sanders, A.t.R.J., Weeghal, I.v., Vogelaar, M., Verheul, W., Pieters, R.H.M., de Wit, N., et al. (2013). Effects of improved patient participation in primary care on health-related outcomes: a systematic review. *Family Practice: The International Journal for Research in Primary Care*, 30, 365-378.

Sarkar, U., Fisher, L., & Schillinger, D. (2006). Is Self-Efficacy Associated with Diabetes Self-Management Across Race/Ethnicity and Health Literacy? *Diabetes Care*, 29, 823-829.

Strecher, V.J., DeVellis, B.M., Becker, M.H., & Rosenstock, I., M. (1986). The Role of Self-Efficacy in Achieving Health Behavior Change. *Health Education Quarterly*, 13, 73-91.

Trachtenberg, F., Dugan, E., & Hall, M.A. (2005). How patients' trust relates to their involvement in medical care: Trust in the medical profession is associated with greater willingness to seek care and follow recommendations. *The Journal of Family Practice*, 54, 344-352.



## 5. Overall Conclusion

In this PhD thesis, I set out to unearth socio-economic inequalities in constructs central to Public Understanding of Science in order to shed light on what might aid us theorise and practice a more democratic form of science and medicine. In **Chapter 2**: I investigated age-related bias in science knowledge; I tried to empirically distinguish between lack of information and misinformation; following the identification of age-related bias, I conducted a sensitivity analysis to check whether using the traditional summed score approach to modelling science knowledge, instead of accounting for measurement non-invariance with relation to age, would lead to substantively different conclusions about genetic research. In **Chapter 3** I introduced a new construct, biomedical self-efficacy, and I examined: knowledge and efficacy gaps between low and high educational groups; I also examined the role of Internet use and paying attention to medical stories online in reducing those gaps. In **Chapter 4** I examined how biomedical efficacy and trust, along with a range of other socio-demographics, affect whether individuals want a say in the decisions regarding their health.

There are three major conclusions derived from **Chapter 2**. First, that ‘don’t know’ answers seem to stand on their own, and could potentially offer a more substantively interesting representation of individuals in terms of their knowledge, or lack thereof, if that is of interest. Moreover, misinformation did not emerge as a separate dimension using this data. Perhaps with another dataset that proposition would hold water. Second, there is very significant measurement non-invariance in relation to age. This effectively means that we may not be able to correctly compare people of different ages in terms of their observed level of knowledge, or lack thereof. That would not take into account the measurement error and heterogeneity of this construct and measurement instrument, and potentially lead to us making erroneous conclusions about what people of different ages know about science. A

separate analysis investigating attitudes to genetics was carried out using two different parameterisations: 1) the latent class framework taking into account that people who say they ‘don’t know’ form a separate group, as well as the measurement non-invariance in relation to age identified in previous analyses; 2) the summed score approach using both a continuous measure as well as a broken down version of it, that is meant to render it comparable to the latent classes. Evidence was found that using the traditional summed scale combining ‘don’t know’ and incorrect answers, whether as a scale or categorical variable, does not introduce bias when examining an attitudinal question such as people’s optimism regarding the potential of genetic research. In other words, applied researchers can rest assured that using the simplest method will most likely not bias their substantive and theoretical conclusions, despite evidence of measurement non-equivalence. This may well be because a misinformed group was not found. In this context, and with this data, admitting that one doesn’t know and trying to provide an answer (that is incorrect) may not trigger different cognitive processes, as misinformation has been found to do. The other possibility is that a misinformed group was not identified because an explicit ‘don’t know’ option was offered. And yet another potential explanation is that ‘probably true/false’ answers were grouped with ‘definitely true/false’ answers because of sample size constraints. Perhaps with different data, a misinformed group of British individuals can be identified, if it really exists.

**Chapter 3** has made three important contributions to the Public Understanding of Science literature, as well as to a wider literature on socio-economic inequalities. First, a new measure, biomedical self-efficacy, was operationalised. Unlike perceived knowledge on its own, this new measure offers additional insight by taking into account how people’s self-reported understanding of key terms in biomedicine works in concordance with their willingness to participate in policy relevant public consultation exercises. The second contribution is that we now have better comprehension of educational inequalities in efficacy

and science knowledge in a UK context. The third, and perhaps most important, output is the finding that the Internet appears to work as both a divider and an equaliser, which is in line with the mixed findings regarding factual and perceived science knowledge. General increased use of the Internet seems to be the most important factor in reducing educational inequities in science knowledge with a focus on biology and genetics. As far as efficacy is concerned, if both those with more education and their counterparts report that they pay attention to medical stories online, inequalities in efficacy augment. Nonetheless, disparities reduce between those lower on education who report remembering seeing information relating to medical research online, and those higher on education that do not report paying attention to online medical research stories. In other words, paying attention will only help the former catch up provided their complements do not report engaging with online content about medical research. The models explain almost a quarter of the variability in efficacy and knowledge.

In the matter of science policy implications, it seems that promoting more biomedical material online can only help reduce the democratic deficit to an extent. Encouragingly, Internet use does promote a better understanding of science knowledge. With regards to their confidence about it, the picture looks slightly different. Using the Internet a lot and paying attention to medical research content only helps reduce the discrepancies very slightly if those highly education do not report engaging with similar content. Having said that, perhaps future studies could employ a better measure of attention and test whether the same holds true. Information dissemination campaigns online seem to have the potential to increase knowledge, but they will not necessarily dwindle the educational differences in involvement in public debates about policy matters. One potential strategy could be making biomedical online content more accessible and even more widely circulated, which may reduce the gap between low educated group who engage with it and highly educated people who don't. The

danger of that would be that the highly educated group who do engage with it would be even further ahead of their counterparts. Further studies should endeavour to break down online activities and examine which types bode better for increasing knowledge and efficacy in the UK. Finally, as explained in the literature review, preferred participation and actual participation do not match, unfortunately. Hence, we cannot draw any particular conclusions about how increased feelings for efficacy might translate into actual participatory behaviour. This is for another study to examine. All we can say is that higher efficacy probably helps more than hinders in getting citizens to engage more.

The research carried out in **Chapter 4** has made three important contributions. The first overall contribution is made to the literature on patient-doctor interactions. Theoretical models of patient-physician interaction have been discussed for decades, but our understanding of where the British public position themselves has so far been relatively poor. Moreover, the scarcity of studies employing representative samples has left a large gap that needs to be addressed. This study has helped shed light on how two polar concepts, namely biomedical self-efficacy and trust in doctors as information sources, work both divergently and in tandem with individuals' socio-economic characteristics in shaping public participation in medical decision-making in a clinical scenario. High efficacy and low trust have been found to lead to more self-reported confidence to challenge their GPs in a clinical scenario. The second contribution related to findings regarding which model of patient-medic interaction the British public seems to prefer. We can think of the levels of strength of confidence to challenge a GP's decisions as indicating adherence to the different types of patient-physician interaction models. Ergo, little or no confidence to challenge a doctor is taken to designate a preference for the paternalistic model. Some confidence is posited to signal a preference for the shared-decision making model. A prominent level of confidence is assumed to indicate preference for the informed model of medical decision-making

Overall, it seems that the British public adhere more to a shared decision-making model. This is supported by the fact that most individuals in this sample show moderate confidence and willingness to participate by saying it is like or somewhat like them to question a GP in a clinical situation. Openness to experience, decreased neuroticism and agreeableness have been found to be related to preferences for a more active decision-making style in other studies, so we can infer that perhaps these personality traits are quite prevalent in the UK. This is encouraging for a democracy, in the sense that people with such personality traits may be easier to convince to be involved in other forms of participation in science. The third contribution addresses a survey methodological gap in the health literature by employing a representative sample of the UK population. Due to the moderate prediction accuracy of the model, no strong policy claims can be made. However, in terms of reducing a democratic deficit in patient participation in modern medical practice, things look quite bright on this front, given that most people seem in favour of a shared decision-making model. Aiming to reduce the gap between men and women, with the latter being more likely to have confidence in a medical situation, is an avenue that could potentially lead to an even fairer medical practice system. Another potential avenue is increasing people's confidence through campaigns that disseminate more attractive, distilled, and widespread information relating to at least various common ailments people present themselves to the doctor with. Future research should explore other factors that might play a part in public attitudes towards participation in medical decision-making to further elucidate this.

**Limitations:** The three empirical chapters presented in this thesis have a number of limitations, which should be borne in mind when considering the findings. Firstly, the response rates to the three Waves of the Wellcome Trust Monitor Survey are quite low, namely: 49.3% for the adult sample at Wave I, 52.6% for the adult sample at Wave II, and

51.4% for the sample at Wave III ( unlike the first two waves, this was not fielded to a young sample).

High response rates normally signify that the data is more likely to be representative of the general population. Survey methodologists generally recommend targets of 70 to 85% response rates (Groves et al., 2009). Having said that, Babbie (2004) suggested that a response rate of 50% is acceptable, and the survey data should not be discarded. A meta-analysis has also recently shown that response rates are not necessarily linked to nonresponse bias (Groves and Peytcheva, 2008). Meterko et al. (2015) also show that ‘low’ response rate surveys may indeed represent attitudes of the general population with accuracy. However, there is no way to test whether respondents who chose to decline to take part in the Wellcome Monitor Survey differ, in terms of their knowledge, efficacy, and attitudes towards participation in medical-decision ( for instance), from those who did. Thus, results from the three empirical chapters should not be interpreted as greatly representative of the general UK population. In other words, any findings are chiefly relatable to this sample only. Policy implications discussed above should thus also be interpreted with caution. However, as Meterko et al. (2015) recommend, these results should not be regarded as uninformative.

Secondly, the Wellcome Trust Monitor sample appears to be very well educated. For instance, in Chapter 3, 50% of the sample at Wave II possesses a higher degree. This is not entirely representative of the population distribution of education in the UK at the time the data was collected. When interviews were conducted in 2012, 35.9% of the UK population possessed a higher education qualification (ONS, 2012). Given that education-based inequalities in knowledge and efficacy are the focus of Chapter 3, the implications are important. Such a well-educated sample is likely to possess more knowledge and have a higher sense of efficacy overall. This can distort our findings related to how great knowledge

and efficacy gaps really are. Put succinctly, the knowledge and efficacy gaps found may be larger than those we may have found with a different sample, more similar, in terms of their educational level, to the UK population.

## References

Babbie E.R. (2004). *The Practice of Social Research*. 13<sup>th</sup> Edition. Belmont, CA: Wadsworth.

Groves, R.M., Fowler, F.J.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. 2<sup>nd</sup> Edition. Hoboken, NJ: John Wiley & Sons.

Groves, R.M., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly*, 72, 2-31.

Meterko, M., Restuccia, J.D., Stolzmann, K., Mohr, D., Brennan, C., Glasgow, J., & Kaboli, P. (2015). *Public Opinion Quarterly*, 79(1), 130-144.

ONS (2012). *Graduates in the Labour Market, 2012*. Office for National Statistics.